Acetylation of intrinsically disordered regions regulates liquid-liquid phase separation

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Makoto Saito

von Tokyo, Japan

Basel, 2019

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Patrick Matthias, Prof. Dr. Markus Affolter

Basel, den 13. November 2018

Prof. Dr. Martin Spiess, Dekan

Summary

Cell signaling networks are regulated by reversible protein post-translational modifications (PTMs). Acetylation of the ε-amino group on lysines was first discovered on histones, and it is now widely accepted as an important modulator for diverse cellular processes. Acetylation is catalyzed by two types of enzymes: lysine acetyltransferases acetylate lysine residues on proteins by transferring an acetyl group from acetyl-coenzyme A, while histone deacetylases (HDACs) remove acetyl groups. Among HDAC family proteins, HDAC6 is a unique cytoplasmic deacetylase with tandem deacetylase domains and a ubiquitin binding zinc-finger domain. HDAC6 has been implicated in several biological processes, such as stress response and regulation of the cytoskeleton. However, there is still little information about how HDAC6 regulates these processes. Moreover, the mechanism how HDAC6 deacetylate its substrates also remains unexplored.

The first, and main part of the thesis is the identification of novel substrates of HDAC6 to clarify how HDAC6 functions in stress response pathway. Previous work from our laboratory had demonstrated that HDAC6 is important for the formation of stress granules (SGs). SGs are membrane-less organelles forming in response to stress, and liquid-liquid phase separation (LLPS) of proteins containing intrinsically disordered regions (IDRs) has been proposed as a mechanism underlying the formation of membrane-less organelles including SGs. We systematically identified HDAC6 substrates using acetylome analysis, and found that the very large majority of the high-confidence HDAC6 target sites map to IDRs. We confirmed that DDX3X, an RNA helicase component of SGs, is a novel substrate of HDAC6. We also revealed that specific stresses elicit activation of the lysine acetyltransferase CBP *in vivo*, leading to acetylation of multiple proteins, including DDX3X. Acetylation of DDX3X-IDR was associated with its inefficient LLPS *in vitro* and the decreased total volume of SGs *in vivo*. Deacetylation of DDX3X by HDAC6 was required for assembly of large, mature SGs. In sum, we define HDAC6 as a global regulator of IDRs, and demonstrated a framework to understand how acetylation/deacetylation of IDRs regulates LLPS spatiotemporally for membrane-less organelle formation *in vivo*.

The second part addresses the mechanistic insights of HDAC6-mediated deacetylation. This work was done in collaboration with Y. Miyake and L. Wang. We solved the crystal structure of both HDAC6 catalytic domains. We proposed a new insight into deacetylation mechanism of its substrate α-tubulin. Furthermore, we provided reasonable explanation for the efficacy of HDAC inhibitors on HDAC6, which form the fundamental basis to develop more potent HDAC6-specific inhibitors in the future.

Therefore, in this thesis, I demonstrate the novel insights about HDAC6-mediated deacetylation and its substrates, and propose a new role of protein acetylation as a regulator of intracellular phase transitions, in particular, in the context of SG formation controlled by HDAC6.

Table of Contents

Introduction

1.1 Lysine acetylation/acylation

1.1.1 Brief history of lysine acetylation/acylation

Cell signaling networks control almost all cellular functions and are regulated by reversible protein post-translational modifications (PTMs). Dozens of different PTMs have been identified so far, including phosphorylation (Pawson and Scott, 2005), acetylation (Choudhary et al., 2014; Verdin and Ott, 2015), methylation (Biggar and Li, 2015; Greer and Shi, 2012), glycosylation (Moremen et al., 2012), and modification by ubiquitin and ubiquitin-like modifiers (Dikic et al., 2009; Swatek and Komander, 2016). Acetylation of the ε-amino group on lysines was discovered on histones by Vincent Allfrey and colleagues in 1964, and they proposed that histone acetylation regulates gene expression (Allfrey et al., 1964). This daring proposal about histone acetylation is now widely confirmed and accepted in the field (Verdin and Ott, 2015). Histone acetylation sites are located primarily in the N-terminal tail, which protrudes from the nucleosome core and provides a signaling platform used to remodel chromatin and regulate transcription. Changes in the charge of the histone tails by acetylation weaken the interaction of histone and DNA. Acetylation also alters histone-histone interactions and their interactions with other regulatory proteins, and provides binding site for specific domains. These changes affect the structure of individual nucleosomes and their folding, leading to permissive chromatin structure for transcription.

It took more than 30 years from the discovery of histone acetylation until the identification of the first acetylation-modifying enzyme (Brownell and Allis, 1995). Numerous studies have subsequently identified three classes of regulatory proteins for acetylation: lysine acetyltransferases (KATs) as ‘writers’, which add acetyl groups to proteins; lysine deacetylases (KDACs) as ‘erasers’, which remove acetyl groups from proteins; and proteins harboring acetyllysine binding domains (including bromodomain (BRD), YEATS domain and double plant homeodomain finger (DPF) domain) as ‘readers’, which interact with acetylated lysines (Kac).

During the past decade, developments in high-sensitivity mass spectrometry analysis have facilitated lysine acetylation research (Choudhary et al., 2014). The advancements can be summarized

in the following three points. First, a large number of acetylation sites in non-histone proteins have been identified. The number is similar to other major PTMs, and indicates the regulatory potential of acetylation at proteomic scale. Second, acetylation has been found in organisms ranging from bacteria to humans, suggesting that some regulatory functions of acetylation are well conserved. Third, this proteomics approach also has recently identified other types of lysine modifications by short fatty acids collectively called acylations: these include crotonylation (Kcr), succinylation (Ksuc), glutarylation (Kglu) among others (Sabari et al., 2017) (Fig. 1).

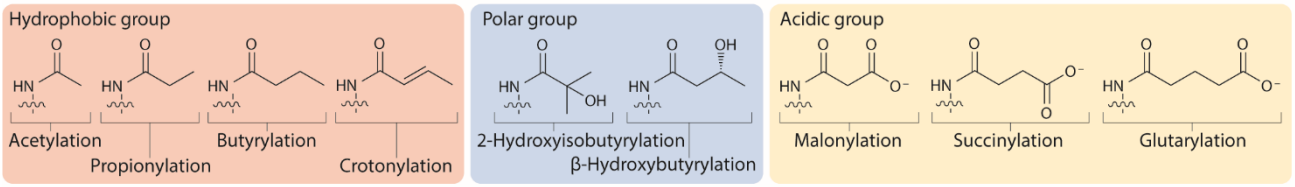


Figure 1. Structures of short-chain acylations on the ε-amine group of Lys, clustered into three groups based on their chemical properties. Adapted from (Sabari et al., 2017).

These findings have broadened the paradigm about lysine acetylation to the acylations, and further nominated these writer, eraser and reader proteins for regulators of protein “acyl”ations. Thus, lysine acetylation and acylation are now widely recognized as general regulators in various cellular functions beyond the initial expected effect on transcriptional regulation.

1.1.2 Writers, Erasers and Readers for lysine acetylation/acylation

Lysine acetyltransferases (KATs)

Lysine acetyltransferases/histone acetyltransferases (KATs/HATs) are enzymes that acetylate lysine residues on proteins by transferring an acetyl group from acetyl-coenzyme A (acetyl-CoA) to form ε-N-acetyllysine. There are three major families of HATs: GNAT (Gcn5-related N-acetyltransferase), MYST (named for the founding members: Moz, Ybf2, Sas2 and Tip60) and the

p300/CREB-binding protein (p300/CBP) families (Table 1). Most HATs form multisubunit complexes, and the combinations of their subunits contribute to the specific characters of each KAT complex. As the major target of these complexes are histones, some subunits have chromatin-binding domains (such as bromodomain, chromodomain and Tudor domain) to recruit the KAT to the appropriate location in the genome. Some KATs can also acetylate non-histone substrates; for example, the tumor suppressor p53 is acetylated by p300/CBP, and this acetylation stimulates its DNA-binding ability (Gu and Roeder, 1997). Besides the function as lysine "acetyl"transferase, KATs belonging to all three major families can also use different acyl-CoAs, which structurally fit their catalytic pockets, as substrates for lysine "acyl"ations. For example, structural studies of p300 suggest that its active site contains a deep aliphatic pocket, which is not seen in other KATs (Kaczmarska et al., 2017). This pocket can accommodate relatively longer acyl-chains, and explains well that p300 is able to catalyze diverse acylations.

Family	HAT	KAT	Histone substrate	Function
GNAT	HAT1	KAT1	H2AK5, H4K5,12	transcriptional activation, DNA repair
	Gcn5	KAT2A	H3K9,14,18,36	
	PCAF	KAT2B	H3K9,14,18,36	
	ELP3	KAT9		
p300/CBP	CBP	KAT3A	H2AK5 H2BK12,15 H3K14,18 H4K5,8	transcriptional activation
	p300	KAT3B	H2AK5 H2BK12,15 H3K14,18 H4K5,8	
MYST	Tip60	KAT5	H4K5,8,12,16	transcriptional activation, DNA repair, replication dosage compensation
	MOZ/MYST3	KAT6A	H3K14	
	MORF/MYST4	KAT6B	H3K14	
	HBO1/MYST2	KAT7	H4K5,8,12	
	HMOF/MYST1	KAT8	H4K16	
TF-related HATs	TAF1/TBP	KAT4		transcriptional activation, Pol III transcription
	TFIIIC90	KAT12	H3 K9,14,18	
NR co-activators	SRC1	KAT13A		transcriptional activation
	AIB1/ACTR/SRC3	KAT13B		
	P160	KAT13C		
	CLOCK	KAT13D		

Table 1. Characteristics of KAT families. Each KAT has unique specificity for each histone variant. Modified figure from (Furdas, 2012; Parthun, 2007).

It is important to note that KATs are not absolutely required for lysine acylation reaction; under specific conditions, some acylation may also occur non-enzymatically. In particular, the higher acyl-CoA concentration and the elevated pH in mitochondria, compared to in the cytoplasm and nucleus, cause the deprotonation of the amino groups and favor non-enzymatic lysine acylation (Wagner and Payne, 2013), and recent proteomics approach for lysine acylation have demonstrated that a large part of the mitochondrial proteins are acylated (discussed in **1.1.3**).

Lysine deacetylases (KDACs)

The mammalian genomes encode 18 enzymes with deacetylase activity (lysine deacetylases/histone deacetylases; KDACs/HDACs), which can be divided into 2 groups: classical HDACs (Fig. 2), whose enzymatic activity requires Zn^{2+}, and sirtuins (Fig. 3), which are NAD^{+}-dependent.

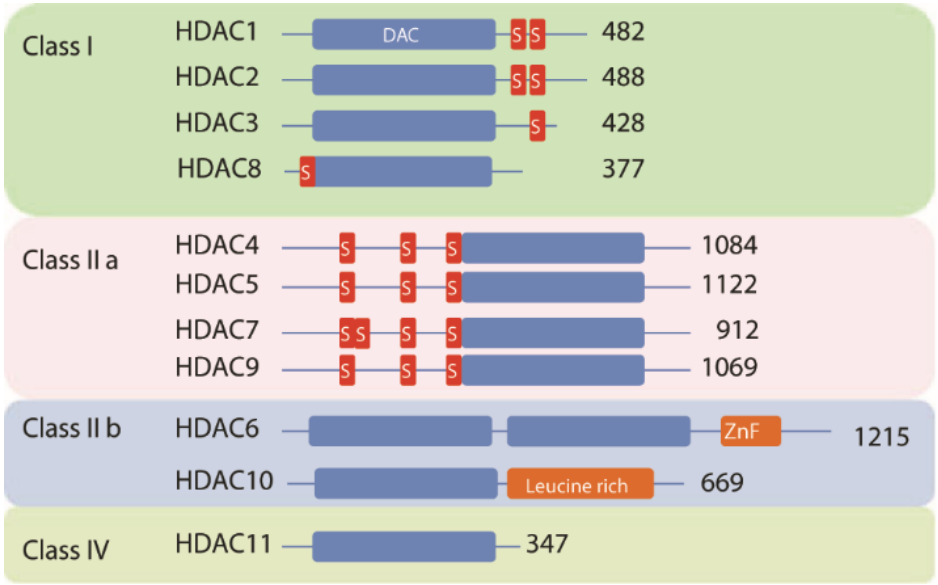


Figure 2. Classical HDAC classification. The domain structures of classical HDACs are shown (numbers refer to amino acid residues in the human proteins). Blue rectangles depict the deacetylase domain, and red rectangles indicate phosphorylation sites. Zinc finger (ZnF) and leucine rich motifs are also indicated. Adapted from (Reichert et al., 2012).

The classical HDAC family harbors 11 members classified into four classes. Class I contains HDAC1, 2, 3, and 8, class IIa comprises HDAC4, 5, 7, and 9, class IIb includes HDAC6 and 10, and class IV comprises HDAC11. The sirtuins are classified as class III HDACs. This family contains seven members (SIRT1-7) which are organized into four classes. SIRT1-SIRT3 belong to class I, SIRT4 to class II, SIRT5 to class III, and SIRT6 and SIRT7 to class IV. The class I HDAC family proteins (HDAC1, 2, 3 and 8) show homology to Rpd3, a founding member from budding yeast, and are ubiquitously expressed. They harbor a conserved deacetylase domain with short N- and C-terminal regions, and they are localized mainly to the nucleus to deacetylate histones. HDAC1 and HDAC2 are generally found together in co-repressor complexes such as the Sin3, NuRD and CoREST complexes. HDAC3 is engaged in a distinct N-CoR complex, and no complexes have been reported for HDAC8. These class I HDACs are localized predominantly to the nucleus to deacetylate histones as their major substrates.

The class IIa HDACs (HDAC4, 5, 7 and 9) have long N-terminal regions where a few serine phosphorylation sites reside. Some kinases, such as calcium/calmodulin-dependent protein kinase (CaMK) and protein kinase D (PKD), phosphorylate these N-terminal regions, allowing 14-3-3 proteins binding and cytoplasmic retention (McKinsey et al., 2000; Vega et al., 2004a; Wang et al., 2000). These regulated phosphorylations of class IIa HDACs provide a mechanism for coupling extracellular signals to transcription and have key roles during development. Unlike other ubiquitously expressed HDACs, class IIa shows restricted expression patterns; HDAC5 and HDAC9 are highly expressed in muscles, heart and brain (Chang et al., 2004; Zhang et al., 2002), HDAC4 is found in the brain and skeleton (Vega et al., 2004b), and HDAC7 is expressed in endothelial cells and thymocytes (Chang et al., 2006b). Class IIa HDACs possess much lower *in vitro* HDAC activity compared to other HDACs, and they have been shown to recruit class I HDACs through their HDAC domains (Fischle et al., 2002); this may contribute to their transcription repression function.

The class IIb HDACs are composed of HDAC6 and HDAC10. HDAC6 is distinct from all other HDACs, as it harbors two deacetylase domains and a C-terminal ubiquitin binding zinc finger

domain. It mainly localizes in the cytosol and contributes to various biological processes such as cytoskeleton control, protein homeostasis and stress response (Matthias et al., 2008) (discussed in **1.2**). HDAC10 has recently been reported as a "PDAC" deacetylating polyamines that are polycations binding to other cellular anionic molecules to stabilize their structure and regulate their functions (Hai et al., 2017); its biological role is still unexplored.

HDAC11, alone class IV member, has small N- and C-terminal regions in addition to its deacetylase domain. Initial study showed that HDAC11 was mainly expressed in brain, heart, muscle, kidney and testis (Gao et al., 2002). However, HDAC11 expression was reported later in hematopoietic cells as well; HDAC11 is bound on the IL10 promoter, and represses its activity in antigen-presenting cells (APCs) (Villagra et al., 2009), suggesting a role in immune functions.

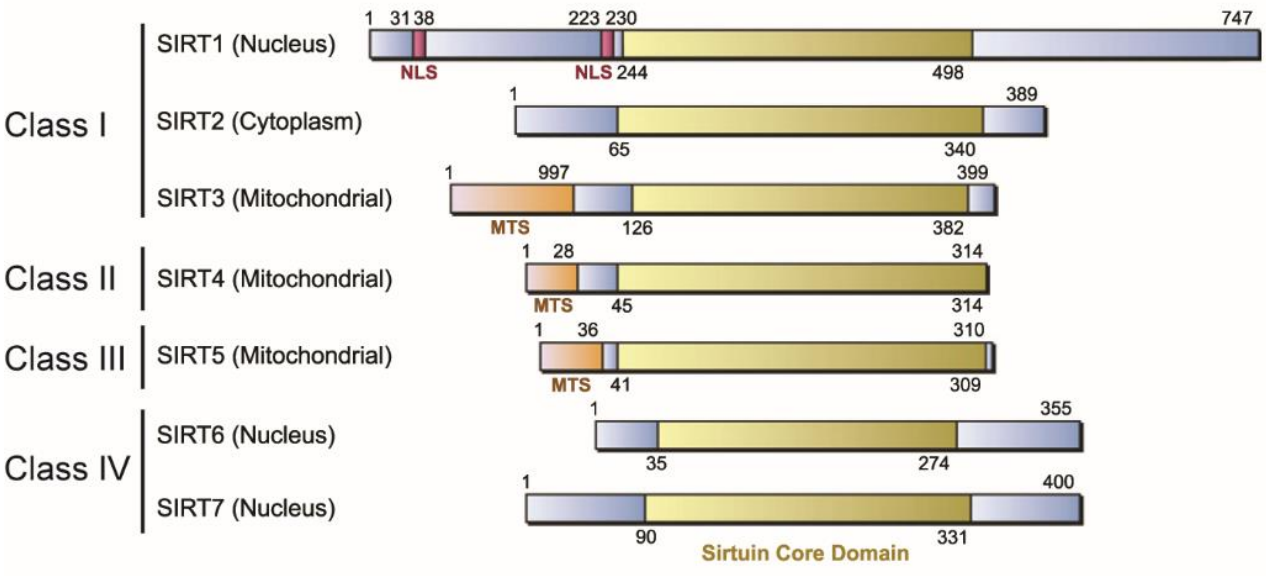


Figure 3. Sirtuin classification. The seven human sirtuins (SIRT1-7) are aligned. Numbers refer to amino acid residues in the human proteins. The conserved core domain that all sirtuins have in common is in yellow. Nuclear localization sequences (NLSs) and mitochondrial targeting sequences (MTSs) are also indicated. Modified figure from (Guarente, 2013).

Class III HDACs, Sirtuins, are NAD^+-dependent deacetylases, which show an individual

pattern of subcellular localization (Fig. 3). Sirtuins deacetylate histones and transcription factors (such as PGC1α and FOXO) in the nucleus, on the other hand, they also deacetylate metabolic enzymes in cytosol and mitochondria (such as PEPCK and IDH2) (Houtkooper et al., 2012). As a result, Sirtuins regulate glucose and lipid metabolism in response to cellular energy levels through deacetylation. Apart from their deacetylase function, some of the members can also work as ADP-ribosyltransferases (SIRT4 and 6), and deacylases (for depropionylation, debutyrylation, decrotonylation, demalonylation, desuccinylation and deglutarylation; SIRT1-5 and 7) (Table 2). Deacylation is likely to be mediated mainly by class III HDACs, although a few studies reported deacylase function of class I HDACs (Wei et al., 2017).

Enzyme class	Lysine acylations							
	Kpr	Kbu	Kcr	Kma	Ksucc	Kglu	Khib	Kbhb
Writers								
p300/CBP	+	+	+	NA	+	+	NA	+
GNATs	+	–/+	–	NA	NA	NA	NA	NA
MYSTs	+	–	–	NA	NA	NA	NA	NA
Erasers								
NAD^+-dependent sirtuins	SIRT1–3	SIRT1–3	SIRT1–3	SIRT5	SIRT3–5,7	SIRT5	NA	NA
Zn^{2+}-dependent HDACs	NA	NA	HDAC3	NA	NA	NA	NA	NA

Table 2. The enzymatic specificities of the different classes of KATs and KDACs with respect to acetylation and acylation. Kbhb, lysine β-hydroxybutyrylation; Kbu, lysine butyrylation; Kcr, lysine crotonylation; Kglu, lysine glutarylation; Khib, lysine 2-hydroxyisobutyrylation; Kma, lysine malonylation; Kpr, lysine propionylation; Ksucc, lysine succinylation; NA, not available. Adapted from (Sabari et al., 2017).

Acetyl-lysine binding proteins

Three domains have been reported to recognize lysine acetylation (Kac). Bromodomains (BRDs) form a conserved structural fold of four α-helices and two inter-helical loops that constitute a hydrophobic pocket. The human proteome encodes 42 BRD-containing proteins, which mainly recognize histone acetylations and regulate gene expression mainly through three mechanisms. First,

some work as scaffolds for other larger protein complexes (e.g. Bromodomain and Extra-Terminal (BET) family proteins). Second, some act as transcription factors or co-regulators (e.g. speckled proteins (SPs)). Third, they have various catalytic functions, including ATP-dependent chromatin remodeling complexes (e.g. SMARCA2/4), methyltransferases (e.g. MLL) and KATs (e.g. PCAF, p300/CBP). Their major targets are histone acetylations, but a few examples demonstrate that some acetyl-lysine binding proteins can recognize non-histone acetylated proteins (Mujtaba et al., 2004; Shi et al., 2014).

Double PHD fingers (DPF) domain is another domain to recognize Kac. Although a large number of PHD-finger containing proteins recognize histone methylation (H3K4me3/H3K9me3/H3K36me3) or unmodified histone tails (H3K4), DPF is highly specific for histone acetylations. So far, DPF domains from five human proteins have been characterized as acetylated lysine readers, including two KATs from the MYST family (MOZ, MORF) and DPF1, DPF2 and DPF3, which are subunits of the BRG1/BRM-associated factor (BAF) chromatin remodeling complex. These proteins use two PHDs to recognize Kac; for example, a PHD of DPF3 binds to the first four N-terminal residues of H3 tail, while another PHD uniquely binds to H3K14Ac in the binding pocket composed of hydrophobic and charged residues (Zeng et al., 2010). This hydrophobic 'dead-end' pocket of PHD1 favors to accommodate lysine crotonylation (Kcr), due to intimate encapsulation and an amide-sensing hydrogen bonding network. (Xiong et al., 2016) (Fig. 4).

The recently identified YEATS (named for Yaf9, ENL, AF9, Taf14 and Sas5) domains also recognize Kac. This domain is also related to transcription regulation (Li et al., 2014). The binding pocket contains an open end where the acetyl group is positioned, suggesting that an extended acyl chain could also be accommodated – recent studies with AF9 have shown that the Kcr is indeed fitting better than Kac in the pocket (Li et al., 2016b). This preferential binding to Kcr is due to the interactions of the crotonylamide group with two sandwiching aromatic residues in the binding pocket, in addition to hydrophobic interactions caused by hydrocarbon extension. In sum, BRDs prefer to bind Kac, while DPF and YEATS domains prefer non-acetyl lysine acylations (Fig. 4).

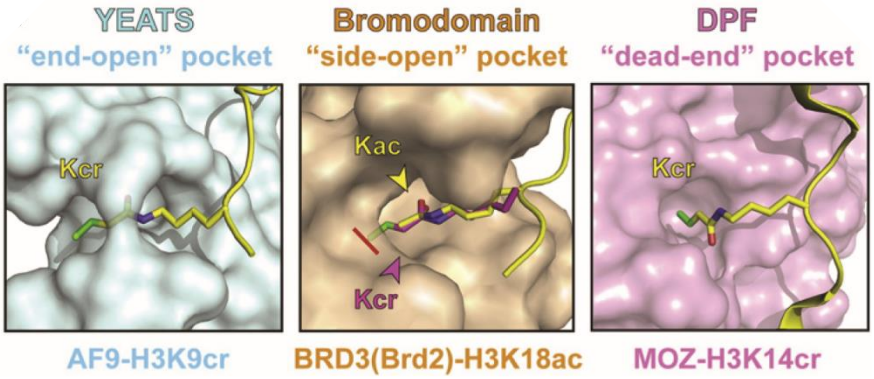


Figure 4. Comparison of the acyl-lysine binding pockets among YEATS (left, PDB: 5HJB), Bromodomain (middle, PDB: 5HJC), and DPF (right, PDB: 5B76). The histone peptide is shown in yellow, with the side chain of acyl-lysine, and the green color in side chain highlights two extended hydrocarbon crotonyl group. Adapted from (Zhao et al., 2017).

1.1.3 Proteomics approach for lysine acetylation/acylation

Recent progress in high-sensitivity mass spectrometry has enabled the identification of acetylation/acylation sites at proteomic scale (acetylome/acylome), which provides important resources to clarify their novel regulatory functions. Although different methodological approaches can be used in the acetylome/acylome analysis, the bottom-up proteomics approach, which involves enzymatic digestion of all proteins in the samples followed by liquid chromatography and tandem MS (LC-MS/MS), is the gold standard in the field (Fig.5). Proteins are extracted from cells, and digested with a protease such as trypsin and Lys-C. The acetylated peptides in the samples are enriched with acetyllysine-specific antibodies. This process is often coupled to sample fractionation in order to reduce complexity and obtain broader dynamic range. The resulting peptides are separated by LC and ionized before entering the mass spectrometer, where mass spectra and fragmentation spectra are measured. The first acetylome analysis identified 388 acetylation sites in 195 proteins among proteins from HeLa cells and mouse liver mitochondria (Kim et al., 2006). In addition to regulators of chromatin, which were well-known targets of acetylation at that time, non-nuclear acetylated proteins with diverse functions were identified. Most strikingly, more than 20% of mitochondrial proteins including many

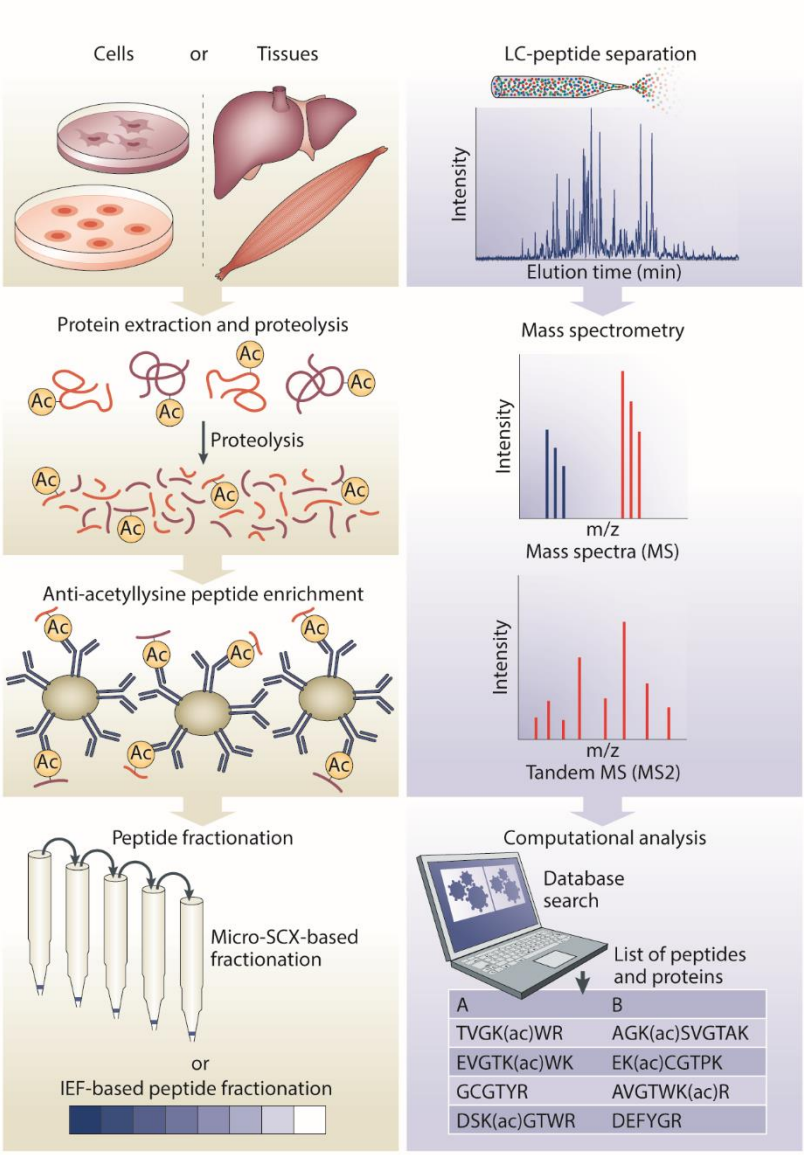


Figure 5. Standard procedure for acetylome analysis. Proteins are extracted from cells or tissues and digested into peptides using proteases. To reduce the sample complexity, acetylated peptides are enriched by immunoaffinity purification with pan-acetyllysine antibodies. The complexity can be further reduced by peptide fractionation methods, such as microscale strong cation exchange chromatography (Micro-SCX) or isoelectric focusing (IEF). After separation by liquid chromatography (LC), peptides undergo electrospray ionization, and peptide ions are transferred into the mass spectrometer. MS and MS/MS spectra are then computationally processed to deduce peptide sequences, including the presence and location of PTMs, and to quantify the abundance of peptides and proteins. Adapted from (Choudhary et al., 2014).

metabolic enzymes were found to be acetylated. Quantitative analyses can identify acetylation sites that are regulated under specific conditions. The most commonly used methods enable the identification by comparing the intensities of acetylated peptides among samples, including metabolic labelling (e.g. Stable Isotope Labelling by Amino acids in Cell culture (SILAC)), chemical labelling (e.g. Tandem Mass Tags (TMT) and isobaric Tags for Relative and Absolute Quantitation (iTRAQ)) and label-free quantification. SILAC has been successfully used to uncover relative changes in acetylome depending on cell types, chemical or genetic perturbations. For example, Choudhary et al. identified 3600 lysine acetylation sites on 1750 proteins from human cell lines, and quantified acetylation changes in response to two deacetylase inhibitors SAHA (suberoylanilide hydroxamic acid) and MS-275 using SILAC (Choudhary et al., 2009). This study revealed that lysine acetylation preferentially targets large protein complexes involved in diverse cellular processes, such as DNA damage repair, RNA splicing, nuclear transport, and cytoskeleton remodeling. The large coverage of the study enabled the authors to analyze structural properties of acetylation; acetylation frequently localizes in structured regions, such as α-helices and β-sheets. This shows stark contrast between acetylation and another PTM, phosphorylation; many phosphorylation sites are found in unstructured regions. At the same time, no clear acetylation motifs have been found, maybe due to the diversity of the KATs/KDACs targeting these sites and also due to the non-enzymatic acetylations. Their experiments were successfully followed by another survey to connect acetylated substrates to the regulator enzymes, using 19 different inhibitors for all 18 human lysine deacetylases (Scholz et al., 2015). Lysine deacetylases inhibitors (KDACis) are used in basic research and also in clinical trials for treatment of cancer and other diseases. However, their specificities in cells still remained to be solved at that time. Over 8,100 acetylation sites in HeLa cells were quantified under KDACis, and the data sets were compared with that of KDAC knockout cells. Another recent study analyzed the CBP/p300-acetylome in time-resolved manner under two catalytic inhibitors and a BRD inhibitor (Weinert et al., 2018). They combined the acetylomes with the transcriptomes which CBP/p300 regulates through acetylation of histones and other transcription regulators, and suggested that the

subset of rapidly regulated CBP/p300-catalyzed acetylation sites is a critical driver of gene transcription.

In parallel with acetylome analysis, proteome analysis for other acylations has also been reported. A study described succinylome from the liver and embryonic fibroblasts of SIRT5 KO mice, where lysine succinylation was increased due to lack of SIRT5 activity (Park et al., 2013). Among the identified 2,565 succinylation sites on 779 proteins, most of the sites (75%) do not overlap with acetylation sites, indicating differential regulation of succinylation and acetylation. These succinylation sites were enriched with enzymes involved in metabolism, including amino acid degradation and TCA cycle. Detailed biochemical analysis revealed that SIRT5 especially repressed two protein complexes, pyruvate dehydrogenase and succinate dehydrogenase. Thus, this work demonstrated widespread roles of lysine succinylation in metabolism. The glutarylome from the liver of SIRT5 KO mice subsequently identified 683 glutarylation sites in 191 proteins (Tan et al., 2014). Glutarylation was also highly enriched in metabolic enzymes, and validated carbamoyl phosphate synthase 1 (CPS1), an enzyme in the urea cycle, as a novel substrate of SIRT5; glutarylation suppressed CPS1 enzymatic activity. These two studies expand the landscape of lysine acylations at proteome scale and increase our understanding of the deacylase SIRT5.

In sum, acetylome/acylome analysis is now widely used in various organisms, leading to the identification of thousands of novel acetylation sites. Such analyses have demonstrated that acetylation sites are frequently conserved across diverse organisms. The most striking observation from acetylome/acylome analysis is that the majority of modified sites are on non-nuclear proteins, including mitochondrial proteome, suggesting an important role for acetylation/acylation in diverse biological functions beyond the control of gene expression.

1.1.4 Functional roles of acetylation/acylation

The best established role of acetylation/acylation is transcriptional regulation. However, it also alters a variety of protein functions by eliminating the positive charge from the ε-amino group of

lysine or introducing steric hindrance, resulting in the modulation of protein interactions with other molecules, changes in enzymatic activity or subcellular localization. Furthermore, it has become clear that acetylation/acylation connects metabolism to cell signaling and gene expression.

Regulation of nucleic acid binding and protein-protein interactions

As discussed above, the first function of acetylation/acylation is regulation of nucleic acid binding ability of proteins. Lysine acetylations on the N-terminal tails of the histones neutralize the positive charge of the tails and decrease their affinity for negatively charged DNA (Hong et al., 1993). This was confirmed well with *in vitro* biochemical experiments, and proposed as a driving force for other epigenetic regulators to access the nucleosome. The second function of acetylation/acylation is regulation of protein-protein interactions. Acetylated lysines can function as docking sites for reader proteins containing BRD, DPF and YEATS domains. Acetylation can also work as a molecular switch to change interaction partner from RNA to proteins. For example, nonacetylated HIV Tat protein preferentially binds to HIV RNA stem-loop structure TAR (transactivating response element) (Mujtaba et al., 2002). Once acetylated by p300, Tat shows increased binding to PCAF, which promotes the association of Tat with RNA polymerase II and HIV gene transcription elongation (Dorr et al., 2002).

Regulation of enzymatic activity

Lysine residues are often present in the active sites of enzymes, and their acetylation/acylation has been implicated in regulating the catalytic activity of various types of enzymes. Activities of KATs themselves are regulated by acetylation; as the autophosphorylation of kinases reflects their activation state, several KATs, including p300/CBP (Thompson et al., 2004) and MYST1 (Yuan et al., 2012), are known to be autoacetylated. Hypoacetylated p300/CBP has a low basal rate of catalytic activity due to autoinhibition by its activation loop, where autoacetylation of several lysines (K1499, K1549, K1554, K1558 and K1560) occurs to enhance its catalytic activity. p300/CBP is autoacetylated as a dimer, suggesting that dimerization triggers autoacetylation and following activation of the enzyme. A large

part of mitochondrial proteins are acylated due to the high concentration of acyl-CoA in the mitochondrial matrix (discussed in **1.1.3**), and their catalytic activities are regulated by these acylations. Acetyl-CoA synthetase 2 (ACECS2) is a mitochondrial matrix enzyme which ligates acetate with CoA to produce acetyl-CoA. ACECS2 is reversibly acetylated at K642 in the active site of the enzyme, and this acetylation inactivates the protein enzymatically. The mitochondrial SIRT3 deacetylates K642 and rescues catalytic activity of ACECS2 (Schwer et al., 2006). Other diverse acylations in mitochondria also tune the function of metabolic enzymes there (Nishida et al., 2015; Park et al., 2013; Tan et al., 2014). Carbamoyl phosphate synthase 1 (CPS1), the rate-limiting enzyme in urea cycle, provides another example (discussed in **1.1.3**) (Tan et al., 2014). There are eight glutarylation sites on CPS1; structure prediction by homology modeling revealed that these sites may be important for binding to either its interaction partner, or CPS1 itself for dimerization, or metabolite as allosteric activator for this enzyme. The mitochondrial SIRT5 removes glutarylation to rescue CPS1 function by these possible mechanisms. In sum, acylation globally regulates catalytic activities of enzymes.

Crosstalk with other PTMs

Acetylation/acylation can compete with other PTMs, such as methylation and ubiquitination, because all these modifications occur on lysine residues. Ubiquitination is a major PTM which regulates protein degradation by proteasome. About one-third of acetylation sites in human cells also undergo ubiquitination (Wagner et al., 2011). Transforming growth factor-β (TGFβ) family members regulate gene expression by receptor-mediated activation of receptor-regulated SMADs (R-SMADs) transcription factors. One of the inhibitory SMADs called SMAD7 forms a complex with a ubiquitin ligase, SMURF1, and inhibits TGFβ signaling by blocking the interaction between the receptor and R-SMADs. In addition, SMURF1 ubiquitinates both the receptors and SMAD7, thereby inducing their degradation. Two ubiquitination sites of SMAD7 by SMURF1 are acetylated by p300, which prevents SMAD7 from degradation (Gronroos et al., 2002). By contrast, the deacetylation of SMAD7 by HDAC1 enhances ubiquitination and degradation (Simonsson et al., 2005). This kind of competitive

crosstalk between acetylation and ubiquitination regulating protein stability is widely observed. In addition to competitive crosstalk, other PTMs can affect acetylation at nearby sites (non-competitive crosstalk). It is well-known that p53 is modified by diverse PTMs, including acetylation, methylation, as well as ubiquitination. In response to DNA damage, the lysine methyltransferase SETD7 methylates p53 at K372 (Ivanov et al., 2007). This methylation stabilizes binding of p53 to chromatin, so that p300/CBP effectively acetylates it at K373 (Dornan et al., 2003). This acetylation also competes with ubiquitination, and stabilized p53 induces subsequent transcription of p21 which promotes cell cycle arrest. This exemplifies that the regulatory mechanism of acetylation also depends on other PTMs.

Regulation of protein localization

Acetylation can regulate the subcellular localization of proteins. The ubiquitin ligase S-phase kinase-associated protein 2 (SKP2) has an important role in tumorigenesis by promoting the destruction of tumor suppressor proteins such as p21 and p27. SKP2 is acetylated by p300 and deacetylated by SIRT3 at K68 and K71, which are located in its nuclear localization sequence (Inuzuka et al., 2012). This acetylation promotes SKP2 nuclear export and, at the same time, inhibits its degradation by a ubiquitin ligase complex APC/Cdh1, resulting in tumor growth. In the same way, subcellular localization of transcription factor FOXO is regulated by crosstalk of phosphorylation and acetylation. In response to insulin, AKT phosphorylates FOXO and induces its nuclear exclusion and subsequent degradation (Brunet et al., 1999). The positive charge of lysines in FOXO1 contributes to its DNA-binding, and acetylation at these residues attenuates its ability to bind target DNA sequence (Matsuzaki et al., 2005). This acetylation is regulated by the counteracting activities of CBP and SIRT1/2 (Frescas et al., 2005; Jing et al., 2007). Phosphorylation of FOXO at S253 by AKT is enhanced by this acetylation, leading to its association with 14-3-3 proteins and its sequestration in the cytoplasm. Thus, acetylation regulates the localization and shuttling between cytoplasm and nucleus of target proteins.

Coupling metabolism with cell signaling and gene expression

Cell signaling and gene expression have been widely known to regulate metabolic processes in response to available nutrients and growth factors with appropriate amounts of metabolic enzymes. In recent years, it has become increasingly clear that cellular metabolites including acyl-CoAs reciprocally regulate signaling networks and gene expression. Acetyl-CoA and other acyl-CoAs, key factors for acetylation/acylation, are also key metabolic intermediates in major carbon catabolic pathways, such as glycolysis and pyruvate oxidation, as well as in β-oxidation of fatty acids. Global histone acetylation levels directly correlate with nuclear acetyl-CoA produced by ATP citrate lyase (ACL) (Wellen et al., 2009). In yeast cells, when cell growth is arrested in the presence of glucose, acetylation is increased specifically in mitochondrial proteins, but not in other compartments, indicating that acetylation in mitochondria is driven by the generation of acetyl-CoA there (Weinert et al., 2014). These two examples clearly demonstrate that the concentration of acetyl-CoA defines the global acetylation level of each cellular compartment. There, acetylation tunes the expression of metabolic enzymes through histone modification and modulate activities of mitochondrial enzymes to adapt cells to specific nutrient conditions. Similar regulatory mechanisms are also likely to take place with other non-acetyl acylations; thus, the concentration of each acyl-CoA species in the compartment may define the global landscape of acylation there (Sabari et al., 2017). This model provides a rationale for some observations. For example, it has been observed that Kcr can be easily detected by mass spectrometry in some tissue-derived cells (Tweedie-Cullen et al., 2012), while enrichment is required to detect it from cells grown in culture media. Cultured cells are usually exposed to higher glucose concentration in cell culture media such as DMEM, than in tissues. Glucose is converted into citrate in mitochondria, and catalyzed by ACL to produce acetyl-CoA. Hence, this artificially enhanced acetyl-CoA may overcome non-acetyl acylations including Kcr in cultured cells. Another example demonstrates that gene expression can be modulated by incorporation of acyl-CoAs into cells. Activation of Toll-like receptor 4 (TLR4) signaling pathway by lipopolysaccharide (LPS) leads to an increase in both Kac and Kcr at loci of p300 target genes. When cells are cultured with crotonyl-CoA

before stimulation by LPS, Kcr level at p300 target genes is increased (Sabari et al., 2015). As positive regulator of transcription AF9 prefers to bind Kcr than Kac (Li et al., 2016b) (discussed in **1.1.2**); this enhanced Kcr efficiently recruits AF9 at p300 target genes, and their transcripiton is also promoted. This kind of acyl-CoA metabolism-mediated transcription plays important role in physiological condition such as ketogenesis. Ketogenesis is a physiological response to low blood glucose and liver glycogen levels; acetyl-CoA is converted to ketone bodies such as β-hydroxybutyrate in the mitochondria of liver cells. The major role of β-hydroxybutyrate is as an energy source, but at the same time, histone Kbhb modification by resulting β-hydroxybutyryl-CoA is subsequently increased. In particular, the expression of genes related to starvation response are upregulated by Kbhb (Xie et al., 2016). Further studies will be required to know how histones are modified by different sets of acyl-CoA depending on their genomic locus and when these mechanisms are essential in physiological conditions other than ketogenesis. Nevertheless, the observations mentioned above clearly demonstrate the importance of acyl-CoA metabolism in cell signaling and gene expression.

1.2 Histone deacetylase 6 (HDAC6)

1.2.1 Domain structure and catalytic activities of HDAC6

Histone deacetylase 6 (HDAC6), which was originally cloned by its homology to the *S. cerevisiae* histone deacetylase HDA1, is one of the class II histone deacetylase (Grozinger et al., 1999; Verdel and Khochbin, 1999). HDAC6 locates mainly in the cytoplasm and has unique structural features (Fig. 6); it contains tandem catalytic domains (CD1 and CD2) and a C-terminal ubiquitin binding zinc finger domain (ZnF-UBP). In addition to these domains, HDAC6 has two nuclear export signals (NESs) and the region called SE14 (only observed in human protein), which ensures to keep HDAC6 in the cytoplasm.

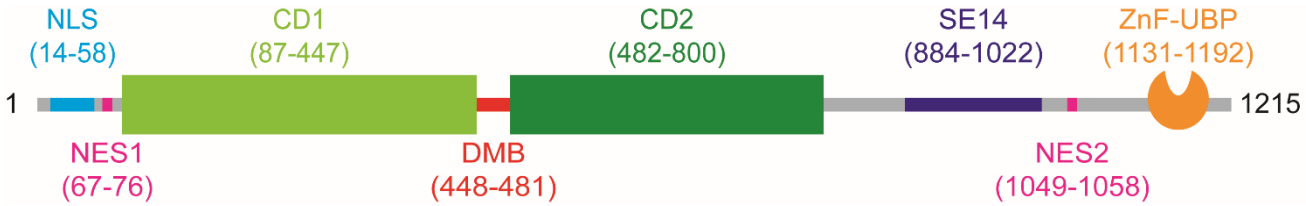


Figure 6. Functional domain organization of human HDAC6 (numbers refer to amino acid residues in the human proteins). NLS: nuclear localization signal. NES1 and 2: nuclear export signal 1 and 2, respectively. CD1 and 2: catalytic domains 1 and 2, respectively. DMB: dynein motor binding domain. SE14 (only in human): anchoring the protein in the cytoplasm. ZnF-UBP: a high affinity ubiquitin-binding motif.

The biological roles of HDAC6 are dependent on these two domains – catalytic domains and ZnF-UBP. HDAC6 primarily functions through altering biochemical properties of its substrates by deacetylation. In addition to its established targets such as α-tubulin, cortactin and heat shock protein 90 (Hsp90) (Hubbert et al., 2002; Kovacs et al., 2005; Zhang et al., 2007; Zhang et al., 2003), several other cytoplasmic proteins have recently been reported as novel substrates. The roles of each substrate will be discussed in the next section (**1.2.2**). HDAC6 ZnF-UBP binds to aggregated proteins and the importance of this function on protein homeostasis will be discussed separately (discussed in **1.2.3**).

1.2.2 HDAC6 substrates

Early work demonstrated that HDAC6 mainly localizes in cytoplasm and functions as tubulin deacetylase. Recent discoveries confirmed the unique functions of HDAC6 among HDAC family proteins, and identified it as a versatile deacetylase involved in wide range of biological processes: HDAC6 regulates cytoskeletal networks, protein folding, stress and viral response, signaling pathways for cell growth, transcription as well as organelle formation through its deacetylase activity. These substrates are listed (Table 3) and summarized schematically (Fig. 7) below, and discussed in more detail in the following section.

Name	HDAC6 species	HDAC6 target sites	Methods to identify / confirm as substrate	Biological meaning of acetylation	Reference
1. α-tubulin	Mouse Human	K40	HD6 KD in NIH3T3 cells.	Mechanical resilience of microtubules	(Hubbert et al., 2002)
2. β-tubulin	Mouse	K58	MS analysis of brain from HD6 KO mice.	Unknown	(Liu et al., 2015)
3. Tau	Human	N.A.	Tubastatin A treatment in OLN-93 cells expressing human Tau.	Protein stability	(Noack et al., 2014)
4. Cortactin	Mouse Human	K124, 161, 235, 309 (Human)	HD6 KD and overexpression in A549 and HeLa cells. MS analysis of 293T. ACY-1215 and Tubastatin A treatment in megakaryocytes.	Localization	(Messaoudi et al., 2017; Zhang et al., 2007)
5. MHC9	Mouse	N.A.	Acetylome analysis of liver from HD6 KO mice.	Binding ability to actin	(Zhang et al., 2015b)
6. Hsp90	Mouse Human	N.A.	HD6 KD in human A431 and A549 cells.	Binding ability to its cochaperone	(Kovacs et al., 2005)
7. Hsp70	Rat	N.A.	HPOB treatment in PC12 cell.	Chaperone activity	(Li et al., 2016c)
8. Hsc70	Mouse	N.A.	Acetylome analysis of liver from HD6 KO mice.	Dimerization ability	(Zhang et al., 2015b)
9. DNAJA1	Mouse	N.A.	Acetylome analysis of liver from HD6 KO mice.	Dimerization ability	(Zhang et al., 2015b)
10. GRP78	Human	N.A.	Panobinostat treatment in MCF7 and MDA-MB-231 cells.	Binding ability	(Rao et al., 2010)
11. Peroxiredoxins	Human	K197	HD6 KD and overexpression in LNCaP and LAPC4 cells, respectively.	Reducing ability	(Parmigiani et al., 2008)
12. TDP-43	Human	K145, 192	HD6 KD with overexpressed TDP-43 in Neuro2A cell.	Aggregates formation	(Cohen et al., 2015)
13. TRIM50	Human	K372	HD6 overexpression in HEK293T cell.	Localization	(Fusco et al., 2014)
14. RIG-I (DDX58)	Mouse Human	K909	HD6 KO MEF cells. HD6 overexpression in HEK293T cells.	Binding ability to RNA	(Choi et al., 2016)

15. Tat	HIV	K28	HD6 KD, overexpression and Tubacin treatment in HEK293T cells.	Transactivator function	(Huo et al., 2011)
16. K-RAS	Human	K104	HD6 KD in SW480 cell.	Binding ability to GEFs	(Yang et al., 2013)
17. ERK1	Human	K72	ACY-1215 treatment, HD6 overexpression in MEFs and HEK293T, A549 cells.	Kinase activity	(Wu et al., 2018)
18. AKT	Human	K163,377	ACY-1215 treatment in human neural progenitor cells.	Kinase activity	(Iaconelli et al., 2017)
19. PTEN	Human	K163	TSA treatment in HEK293T cells.	Localization	(Meng et al., 2016)
20. MST1	Not mentioned	K35	HD6 overexpression of HDAC6 and MST1 in HEK293T cells.	Protein stability	(Li et al., 2016a)
21. Pin1	Human	K46	HD6 overexpression in MCF7 and MDA-MB-231 cells.	Catalytic activity	(Nogues et al., 2016)
22. 14-3-3ζ	Human	K49, 120	Identified by acetylome. Confirmed by Tubacin treatment in MDA-MB-231 cells.	Binding ability to its client proteins	(Mortenson et al., 2015)
23. β-catenin	Human	K49	HD6 KD in HCT116 cells.	Localization	(Li et al., 2008)
24. p53	Human	K120 K381,382	K120: ACY-1215 treatment in TOV-21G cells. K381, 382: HD6 KD in HCT116 and HT29 cells.	Binding ability to a pro-apoptotic protein (K120) Transcription activity (K381 and 382)	(Bitler et al., 2017) (Ryu et al., 2017a)
25. RelA/p65	Human	N.A.	Tubastatin A treatment in NU-DUL-1 cells.	Localization and transcription activity	(Jia et al., 2017)
26. HIF1 α	Mouse	N.A.	HD6 KO MEF cells.	Protein stability	(Ryu et al., 2017b)
27. HMGN2	Human	K2	Bufexamac treatment in MCF7 cells.	Co-activator function	(Medler et al., 2016)
28. Ku70	Human	K539, 542	Tubacin treatment in SH-SY5Y cells.	Binding ability to a pro-apoptotic protein	(Subramanian et al., 2011)
29. Survivin	Not mentioned	K129	Overexpression of HDAC6 and MST1 in MCF7 cells.	Dimerization ability	(Riolo et al., 2012)
30. MSH2	Mouse Human	K845, 847, 871, 892 (Human)	HD6 KO MEF cells. MS analysis of overexpressed MSH2 in 293T under TSA.	Protein stability	(Zhang et al., 2014)

31. Mfn1	Mouse	K222	HD6 KO MEF cells.	Mitochondrial fusion	(Lee et al., 2014a)
32. CIDEC	Mouse	K56	Acetylome analysis of HD6 KO mice.	Protein stability	(Qian et al., 2017)
33. hERG	Human	N.A.	Overexpression of HDAC6 and hERG in HEK293T cells.	Protein stability	(Li et al., 2018)
00. DDX3X	Mouse	K118	Acetylome analysis of MEFs and HeLa cells.	Phase separation and stress granule formation	(Saito et al., 2019)

Table 3. Summary of HDAC6 substrates. It is important to note that 00. DDX3X was identified by the work for this thesis, and it is separately discussed in detail (**2.1**).

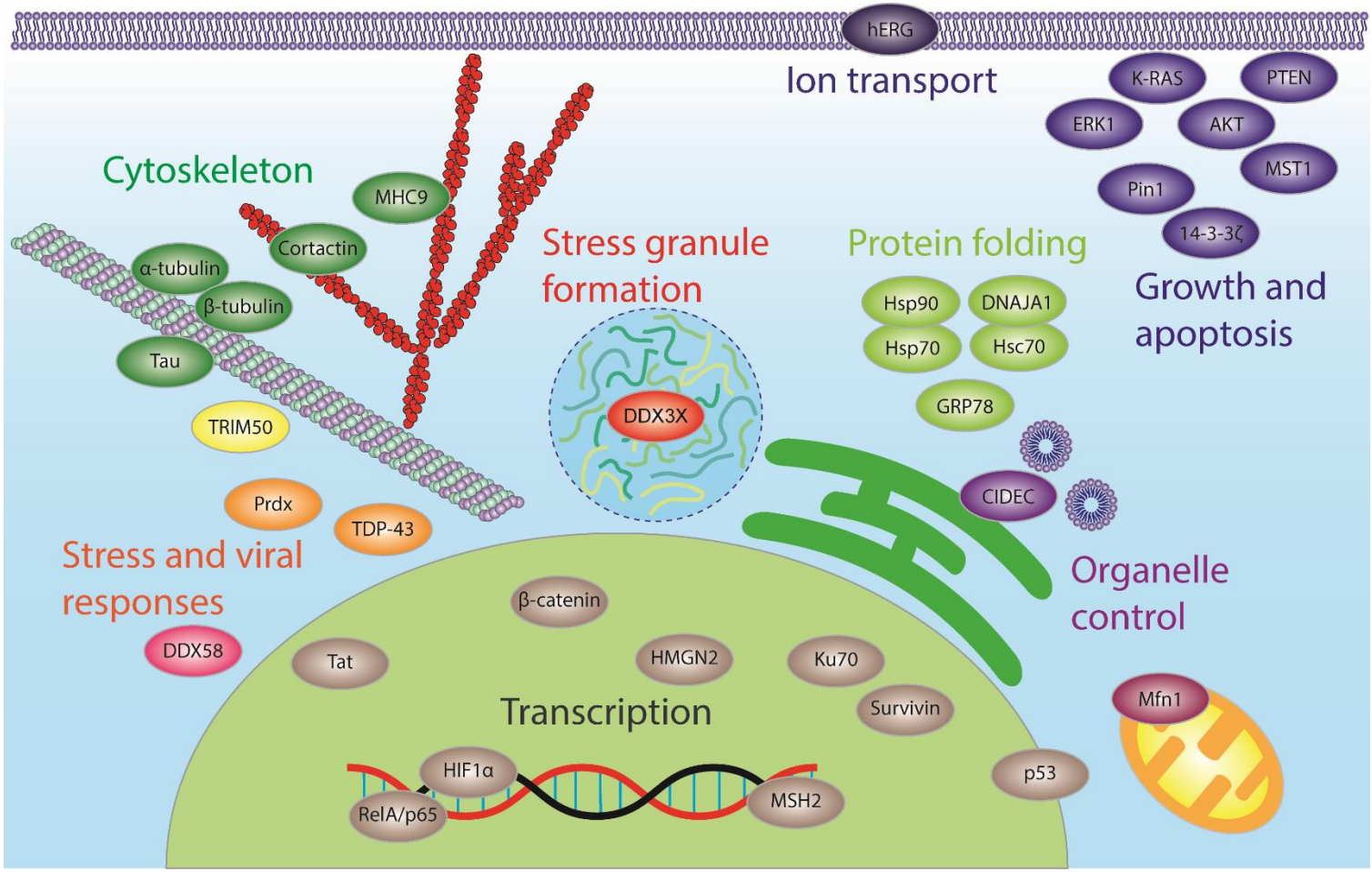


Figure 7. Diverse functions of HDAC6 substrates. Reported 33 HDAC6 substrates are divided into eight groups: cytoskeleton, stress and viral responses, protein folding, organelle control, growth and apoptosis, ion transport, transcription and stress granule formation. It is important to note that stress granule formation through DDX3X deacetylation was revealed by the work for this thesis, and it is separately discussed in detail (**2.1**).

Regulation of cytoskeleton: α/β-tubulins, Tau, cortactin, and MHC9

The major biological function of HDAC6 is regulation of cytoskeletal networks. HDAC6 regulates microtubules through deacetylation of its components, tubulins with associated protein Tau, and it also regulates actin filaments through deacetylation of cortactin and Myosin heavy chain 9 (MHC9) as actin associated proteins.

The first reported *in vivo* HDAC6 substrate is α-tubulin at K40 (Hubbert et al., 2002; Zhang et al., 2003). α-tubulin acetyltransferase (α-TAT) prefers to acetylate K40 of α-tubulin on microtubules over α/β-tubulin dimers (Maruta et al., 1986) and this preference was also confirmed for HDAC6 initially (Hubbert et al., 2002). However, recent studies including ours challenge this; HDAC6 can deacetylate α-tubulin K40 in the lumen of microtubules, but its preferred substrate is unpolymerized tubulin (Miyake et al., 2016; Skultetyova et al., 2017). Although the biological meaning of K40 acetylation has been a long-standing question, recent studies suggested that acetylation increases mechanical resilience to protect long-lived microtubules from mechanical breakage (Portran et al., 2017; Xu et al., 2017). Another microtubule component, β-tubulin has also been reported as a potential HDAC6 substrate by quantitative proteomic analyses (Choudhary et al., 2009; Liu et al., 2015). Importantly, these proteomic analyses also detected additional acetylation sites on α-tubulin other than K40. However, it is still unexplored how these acetylations affect the function of microtubules. HDAC6 also deacetylates the microtubule-associated protein Tau which has an important role in microtubule assembly and stabilization. Acetylation of Tau at several lysine residues tunes its stability, and these acetylations are removed by HDAC6 (Min et al., 2010; Noack et al., 2014).

The formation of actin filaments is also regulated through HDAC6-mediated deacetylation of several effector proteins including cortactin and MHC9. An actin binding protein, cortactin interacts with actin filaments to increase cell motility (Bryce et al., 2005). The positive charge of lysine residues in its repeat region promotes the actin binding ability of cortactin, which is attenuated by acetylations in the region. HDAC6 leads to actin filament remodeling through cortactin deacetylation (Zhang et al., 2007), which is important for angiogenesis in developmental process (Kaluza et al., 2011), platelet

production (Messaoudi et al., 2017) and quality-control autophagy (Lee et al., 2010) (discussed in **1.2.3**). On the other hand, actin binding ability of MHC9 is increased by its acetylation (Zhang et al., 2015b). Thus, inhibition of HDAC6 induces distribution of MYH9 to the actin filaments at cell-cell junctions, which also may explain the effect of HDAC6 inhibition on cell movement.

Protein folding: Hsp90, Hsp70, Hsc70, DNAJA1 and GRP78

HDAC6 is involved in protein quality control by modifying the activity of some chaperone proteins. Heat shock protein 90 (Hsp90) becomes acetylated, and this acetylation prevents it from binding to its cochaperone p23 and impairs the chaperone activity, which is rescued by HDAC6 (Kovacs et al., 2005). Glucocorticoid receptor (GR), a Hsp90 clients, becomes defective in ligand binding, nuclear translocation and gene activation in the absence of HDAC6. Another chaperone protein Hsp70 is also involved in regulation of GR by HDAC6. (Li et al., 2016c). Heat shock factor 1 (HSF1) in $Foxp3^+$ T-regulatory cells (Tregs) represents another example in HDAC6-Hsp90 regulatory axis (de Zoeten et al., 2011). The chaperone proteins Hsc70 and DNAJA1 have recently been identified also as HDAC6 substrates (Zhang et al., 2015b). Deacetylation of these chaperone proteins by HDAC6 helps their dimerization, although the biological importance is still unclear. Glucose-regulated protein 78 (GRP78, also called BiP, HSPA5) is an endoplasmic reticulum (ER) specific homolog of Hsp70 (Ron and Walter, 2007). In stressed condition, accumulated misfolded proteins in ER bind to GRP78 and induce its dissociation from the receptor, leading to activation of Protein kinase RNA-like ER kinase (PERK) for the following stress response. HDAC6 also localizes in the ER and deacetylate GRP78 (Rao et al., 2010), which promotes GRP78 binding to PERK. Thus, HDAC6 regulates a variety of chaperone proteins by changing the binding ability to their interaction partners, including client proteins, through deacetylation. In each case, however, it is still unknown how large the population of client proteins affected by HDAC6-mediated deacetylation is. Apart from tuning chaperone activity by deacetylation, HDAC6 directly interacts with another chaperone protein, VCP/p97, and controls the fate of ubiquitinated misfolded proteins (Boyault et al., 2007). However, VCP/p97 itself has not

been reported as a substrate (discussed in **1.2.3**).

Stress and viral response: Peroxiredoxins, TDP-43, TRIM50, RIG-I and Tat

Cellular responses to stresses, including oxidative stress, protein misfolding, and even viral infection, is also mediated by HDAC6. HDAC6 regulates antioxidant proteins to cope with oxidative stress, and also prevents protein aggregations in oxidative stress. Peroxiredoxins (Prx) family proteins function as antioxidants; they reduce elevated H_2O_2 level through oxidization of their cysteine residue. Acetylation increases this reducing activity, which is counteracted by HDAC6 (Parmigiani et al., 2008). As antioxidant activity of Prx is linked to resistance to chemotherapy and radiation therapy for cancer (Chen et al., 2002), inactivating Prx would be beneficial in HDAC6-targeted cancer therapy. An RNA-binding protein, TDP-43 is another HDAC6 substrate associated with oxidative stress response. TDP-43 is recurrently mutated in inherited form of the neurodegenerative disease such as amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration with ubiquitin-positive inclusions (FTLD-U) (Lagier-Tourenne and Cleveland, 2009), and its aggregates are observed in the cytoplasm and nucleus of neurons in the patients (Neumann et al., 2006). This aggregation is likely to be enhanced by acetylation; immunohistochemistry with cortical brain and spinal cord sections of the patients shows that TDP-43 in pathological aggregates is acetylated. Acetylations occur in RNA-binding motif of TDP-43 (K145 and K192), and weaken its RNA-binding function. Oxidative stress promotes TDP-43 acetylation, and HDAC6 is required to deacetylate TDP-43 and prevent its aggregation (Cohen et al., 2015).

Apart from TDP-43, HDAC6 is generally important for aggregate formation, which are often associated with neurodegenerative diseases (discussed in **1.2.3**). Aggregated proteins, which are not efficiently degraded by ubiquitin-proteasome system (UPS), often accumulate into microscopically visible structures called as aggresomes (Kopito, 2000). HDAC6 binds to ubiquitinated and aggregated proteins through its ZnF-UBP domain, and directly brings them into an aggresome using dynein motor system (Kawaguchi et al., 2003). Moreover, HDAC6 also deacetylates K372 of TRIM50, which is

important for aggresome clearance, to enhance its localization into aggresome (Fusco et al., 2012). As mentioned below (in **1.2.3**), influenza virus A (IAV) hijacks this HDAC6-mediated aggresome formation pathway to be efficiently uncoated in host cells (Banerjee et al., 2014), indicating the role of HDAC6 in viral infection. In parallel, HDAC6 also regulates infectivity of other RNA viruses through deacetylation of RIG-I (Retinoic-acid-inducible gene-I, also known as DDX58). RIG-I functions as a sensor of viral infection; viral RNAs induce conformational change of RIG-I so that it can interact with mitochondrial antiviral signaling proteins to induce innate immunity responses such as the production of type I interferons (Leung and Amarasinghe, 2012). Acetylations of RIG-I at K858 and K909 impair its binding to dsRNA. In response to RNA viral infection, HDAC6 interacts with and deacetylates RIG-I, thereby increasing its viral RNA sensing ability (Choi et al., 2016). Accordingly, HDAC6 KO mice are more susceptible to an RNA virus (vesicular stomatitis virus) (Choi et al., 2016). Furthermore, HDAC6 also deacetylates viral protein itself; K28 of HIV-1 Tat is acetylated by p300/PCAF, while it is deacetylated by HDAC6 (Huo et al., 2011). This deacetylation impairs transactivation activity of Tat toward HIV-1 transcription. These findings indicate the multifaceted roles of HDAC6 in RNA virus infection.

Growth and apoptosis signaling pathways: K-RAS, ERK1, AKT, PTEN, MST1, Pin1 and 14-3-3ζ

Several components of phosphorylation signaling, including the MAPK, PI3K-AKT and Hippo pathways, have been reported as HDAC6 substrates. Upon stimulation of growth factors, RAS GTPase protein becomes its active form to activate downstream signaling proteins. RAS proteins cycle between ‘active’ and ‘inactive’ states that are conferred by the binding of GTP and GDP (Pylayeva-Gupta et al., 2011), and the GTP-bound form activates several downstream signaling pathways. One of RAS isoforms, K-RAS, becomes acetylated at K104 (Yang et al., 2012), and this acetylated form is resistant to guanine nucleotide exchange factors (GEFs)-mediated activation; thus acetylated K-RAS cannot maintain a constitutively activated state. This K104 acetylation of K-RAS is removed by

HDAC6 and SIRT2, resulting in cancer cell survival (Yang et al., 2013). Two representative downstream kinases are also targeted by HDAC6; ERK1 (Wu et al., 2018) and AKT (Iaconelli et al., 2017) are activated by HDAC6-mediated deacetylation. The acetylation site of ERK1 (K72) regulated by HDAC6 is located near the ATP binding site, and stabilizes one wall of the ATP binding site via intramolecular contacts. Acetylation reduces enzymatic activity of ERK1 through the conformation change of the ATP binding site, which is recovered by HDAC6. Two acetylation sites in the kinase domain of AKT, K163 and K377, are deacetylated by HDAC6 (Iaconelli et al., 2017). As is the case of ERK1, K163 of AKT is located in the ATP binding site, and this acetylation decreases kinase activity of AKT. As AKT activity is important during neural progenitor cells (NPCs) differentiation (Zhang et al., 2010), HDAC6 inhibition in NPCs impedes their differentiation into neurons, and promotes the glial lineage. A negative regulator of PI3K-AKT pathway, the tumor suppressor PTEN is impaired by HDAC6-mediated deacetylation. PTEN dephosphorylate PIP_3, thereby attenuating PI3K-AKT signaling (Song et al., 2012). Acetylation of PTEN at K163 inhibits the intramolecular interaction between C-terminal tail of PTEN and its remaining region, which brings PTEN to the cell membrane for its catalytic function. K163 deacetylation is mediated by HDAC6; inhibition of HDAC6 induces PTEN acetylation and subsequent membrane translocation to enhance its tumor suppressor activity (Meng et al., 2016). Another tumor suppressor, MST1, is also negatively regulated by HDAC6. MST1 inhibits the activation of YAP/TAZ, key transcription co-activators in Hippo pathway to prevent the expression of pro-survival target genes. HBXIP, which is upregulated in breast cancer (Liu et al., 2014), induces expression of HDAC6, which deacetylates K35 of MST1. This deacetylation of MST1 promotes its interaction with Hsc70, resulting in lysosome-dependent degradation through chaperone-mediated autophagy (Li et al., 2016a).

Regulatory proteins for serine/threonine kinase signaling pathways mentioned so far have also been reported as HDAC6 substrates. Pin1 is a peptidyl-prolyl isomerase (PPIase); it specifically recognizes phosphorylated Pro-directed Ser/Thr (pSer/Thr-Pro) sequence, which is often catalyzed by MAPKs, and adopts its cis and trans conformations (Lu and Zhou, 2007). Pin1 plays a central role in

tumor progression by activating oncoproteins, and inactivating tumor suppressor proteins by its PPIase function (Zhou and Lu, 2016). In breast cancer cell lines, HDAC6 becomes phosphorylated by G-protein-coupled receptor kinase 2 (GRK2) and induces deacetylation of Pin1 (Nogues et al., 2016). GRK2/HDAC6-mediated deacetylation enhances Pin1 stability, and also increases its binding to mitotic cellular substrates. In addition, regulatory 14-3-3 proteins, which are phospho-serine/threonine binding proteins, are also HDAC6 substrates (Morrison, 2009). An isoform 14-3-3ζ suppresses apoptosis via interaction with pro-apoptotic Bad (Datta et al., 2000). Acetylations at two lysine residues located in the binding pocket (K49 and K120) impair the interaction between 14-3-3ζ and its interaction partners including Bad, which is counteracted by HDAC6 to protect cells from apoptosis (Mortenson et al., 2015).

Nuclear proteins: β-catenin, p53, RelA/p65, HIF1α, HMGN2, Ku70, survivin and MSH2

Some nuclear proteins are deacetylated by HDAC6 when they localize in the cytosol. Furthermore, some reports suggest that HDAC6 can localize in the nucleus and interact with its substrates there. β-catenin, a key transcription factor in Wnt signaling pathway (Clevers and Nusse, 2012), is also regulated by EGF signaling; it releases β-catenin from E-cadherin complex and enhances nuclear localization of β-catenin (Lu et al., 2003). Upon EGF stimulation, HDAC6 is recruited to caveolae and deacetylates β-catenin at K49, leading to inhibition of β-catenin S45 phosphorylation and its nuclear localization (Li et al., 2008). p53 is another transcription factor targeted by HDAC6. Mitochondrial p53 acetylated at K120 promotes apoptosis through decreasing mitochondrial membrane potential (Chen et al., 2011). Cancer cells with ARID1A mutation exhibit higher HDAC6 expression, and HDAC6 inactivates the apoptosis-promoting function of p53 through deacetylation of K120 (Bitler et al., 2017). Apart from mitochondrial function, acetylation of p53 is generally known to be important for its transcriptional activator functions (Gu and Roeder, 1997) due to its effect on DNA binding. In agreement with this, K381/382 of p53 are deacetylated by HDAC6, resulting in impaired transcriptional activation by p53 (Ryu et al., 2017a; Ryu et al., 2018). Two other transcription

factors are also reported as HDAC6 substrates: RelA/p65 (Jia et al., 2017) and HIF1α (Ryu et al., 2017b). RelA/p65 is a negative regulator of miR-27b expression in diffuse large B-cell lymphoma (DLBCL) cells. Acetylation caused by HDAC6 inhibition impairs nuclear localization of RelA/p65, leading to elevated miR-27b expression. As miR-27b represses the expression of the c-Met receptor tyrosine kinase, which is important for lymphoma cell survival, HDAC6 inhibition may be an attractive therapeutic strategy in DLBCL (Jia et al., 2017). HIF1α, an essential transcription factor for cellular response to hypoxia, becomes acetylated by HDAC6 inhibition, and this acetylation promotes HIF1α degradation under hypoxia (Ryu et al., 2017b), suggesting that a role of HDAC6 in the hypoxia response pathway is to regulate the level of HIF1α. In addition to transcription factors, HDAC6 also deacetylates the chromatin remodeling protein HMGN2 to regulate activity of a transcription factor Stat5a (Medler et al., 2016).

HDAC6 regulates apoptosis through deacetylation of nuclear proteins. Ku70 is a nuclear protein which binds to Bax, a pro-apoptotic Bcl2 family protein, to prevent localization of Bax on mitochondria. HDAC6 deacetylates Ku70 to keep the Ku70-Bax complex in the cytosol; in agreement with this, HDAC6-specific inhibition triggers Ku70 acetylation and cell death (Subramanian et al., 2011). Survivin is another anti-apoptotic protein regulated by HDAC6. K129 of survivin is acetylated by CBP, which promotes homodimerization of survivin and maintains its nuclear localization. HDAC6 localizes to nucleus to deacetylate survivin in the estrogen-receptor-positive breast cancer cell line MCF7, and induces survivin to localize in cytosol for its anti-apoptotic activity (Riolo et al., 2012).

HDAC6 is also involved in DNA repair through deacetylation of MutS protein homolog 2 (MSH2). MSH2 is a component of the DNA mismatch repair (MMR) system and works as a heterodimer with MSH6 (called as MutSα) or MSH3 (MutSβ) (Li, 2008). As acetylation sites (K845, 847, 871, and 892) are also targeted by ubiquitination, HDAC6 leads to ubiquitination-dependent degradation of MSH2. Instability of MSH2 induced by HDAC6 decreases the amount of MutSα complex, but not MutSβ, and causes DNA mismatch repair activity (Zhang et al., 2014).

Other biological functions: Mfn1, CIDEC, and hERG

Formation of some organelles such as mitochondria and lipid-droplets (LDs) are impacted by HDAC6. Mitochondrial fusion is a two-step process, where outer membrane fusion is regulated by Mfn1 and Mfn2, while inner membrane fusion is regulated by Opa1. Deacetylation of Mfn1 by HDAC6 is required for mitochondrial fusion (Lee et al., 2014a); mitochondria in HDAC6 KO MEFs do not undergo fusion and are fragmented in glucose-starvation. The location of the acetylation site K222 in the dynamin-related GTPase domain of Mfn1 suggests that acetylation may inhibit its GTPase activity for mitochondrial fusion although this has not been confirmed yet. LDs are tissue-specific organelles, which regulate the storage and hydrolysis of neutral lipids, and are found largely in the adipose tissue. LD formation is regulated negatively by HDAC6 through deacetylation of CIDEC (Qian et al., 2017). It is one of the LD-associated proteins that promote LD fusion in adipocytes (Gong et al., 2009). Acetylation of CIDEC at K56 by PCAF is caused by fatty acids, and leads to a conformational change which prevents ubiquitination and degradation. Therefore, deacetylation of CIDEC by HDAC6 results in its decreased stability, which impairs LD fusion and growth.

In addition to organelle formation, regulation of ion channel is likely to be regulated by HDAC6. The human ether-a-go-go-related (hERG) gene encodes the α subunit of a rapidly activating delayed-rectifier potassium (IKr) channel, whose mutations cause long QT syndrome type 2 (LQT2). HDAC6 directly interacts and deacetylates hERG, resulting in its increased ubiquitination and degradation (Li et al., 2018). Thus, the repertoire of HDAC6 substrate is still expanding, and more comprehensive studies will be required.

1.2.3 HDAC6 as a unique regulator of protein homeostasis

The functional uniqueness of HDAC6 among HDAC family proteins is mainly due to its involvement in protein homeostasis. Misfolded proteins are often polyubiquitinated and degraded by proteasome, a process called ubiquitin-proteasome system (UPS). However, the UPS does not work efficiently for degradation of aggregated proteins and it is known that polymerization of aggregated

proteins can form microscopically visible structures such as inclusion bodies and aggresomes (Kopito, 2000). Inclusion bodies are multiple intracellular foci where misfolded proteins are sequestered, while the aggresome is a special type of inclusion bodies that is usually found at the microtubule organizing center (MTOC). The aggresome was first identified in the study of a mutant of the cystic fibrosis transmembrane conducting regulator CFTR-ΔF508, which tends to be misfolded (Jensen et al., 1995). Under normal conditions, misfolded CFTR-ΔF508 is degraded by the proteasome. However, upon proteasome inhibition, CFTR-ΔF508 forms aggregates and concentrates into a juxtanuclear inclusion body that was named aggresome (Johnston et al., 1998). Further studies revealed that other misfolded proteins also form an aggresome which is cleared by subsequent autophagy. In addition to aggresome, more transient and dynamic condensates with proteins and RNAs, called cellular granules (also known as membrane-less organelles or biomolecular condensates) (discussed in **1.3**) are also observed in cells. Stress granules (SGs) are an example of cellular granules composed of ribonucleoproteins (RNPs) which form when translation initiation is impaired; they are thought to represent a pool of mRNAs stalled in translation (discussed in **1.3.3**). SGs are also cleared by autophagy (Buchan et al., 2013) as well as aggresome mentioned above. As mentioned below, aggresome and SGs seem to be functionally interconnected to each other although the molecular mechanisms which explain the transition of these two states are still missing. HDAC6 impacts several steps of these protein homeostasis processes: UPS, aggresome formation, SGs formation and their clearance by autophagy through six mechanisms discussed below and schematized in Fig. 8.

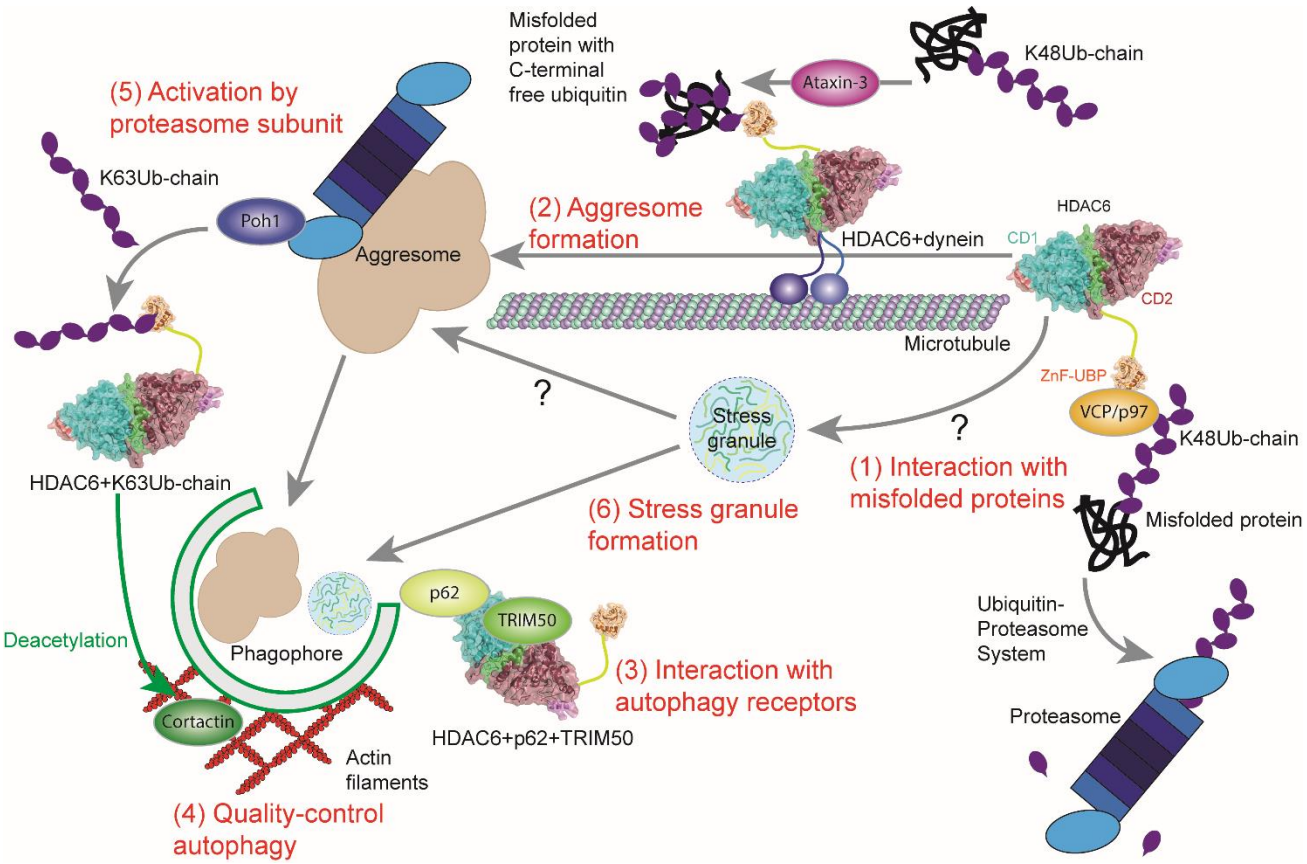


Figure 8. HDAC6 in protein homeostasis. (1) Unfolded/misfolded proteins are usually cleared by ubiquitin-proteasome system (UPS). VCP/p97 removes ubiquitinated proteins from HDAC6 complex to enhance their processing by proteasome. (2) If unfolded/misfolded proteins are not degraded by the proteasome appropriately, they may form aggregates. These aggregates are processed by ataxin-3 and transported by HDAC6/dynein complex to aggresome. (3) Another complex, HDAC6 with TRIM50 and an autophagy receptor p62 also brings aggregated proteins into the aggresome. (4) HDAC6 regulates quality-control autophagy by actin filament remodeling through deacetylation of cortactin to clear aggresomes and stress granules (SGs). (5) Poh1, a proteasomal deubiquitinase produces C-terminal free unanchored K63 ubiquitin chains in aggresome, resulting in HDAC6 activation there and subsequent quality-control autophagy. (6) HDAC6 regulates SG formation, although the molecular mechanism is not well explored when this thesis was started.

(1) HDAC6 uniquely has a ZnF-UBP domain and binds to C-terminal diglycine motif of unanchored ubiquitin (Boyault et al., 2007; Ouyang et al., 2012). This makes HDAC6 an essential factor to coordinate the responses to cytotoxic ubiquitinated misfolded proteins. It binds ubiquitinated

proteins with high-affinity to be hindered from the recognition by ubiquitin specific proteases (USPs) and proteasomes while an HDAC6 interaction partner, VCP/p97 chaperone protein removes them from HDAC6 to enhance their processing by USPs or proteasome (Boyault et al., 2006). This suggests that HDAC6 inhibits misfolded protein degradation by UPS, and that cellular concentrations of these two proteins strongly affect the fate of ubiquitinated proteins. HDAC6 thus prevents ubiquitinated proteins from being delivered to the proteasome.

(2) HDAC6 also works as an adaptor for the dynein motor protein system (Kawaguchi et al., 2003). Upon proteasome inhibition, HDAC6-enriched inclusion body appears and this is indeed aggresome based on its localization and protein components. HDAC6 binds dynein through its binding motif sequence between two catalytic domains, and proteasome inhibition stimulates HDAC6-dynein-ubiquitinated protein complex formation. HDAC6 directly brings ubiquitinated misfolded proteins into an aggresome using the dynein motor system. It is important to note that, from a structural point of view, HDAC6 cannot recognize polyubiquitin chains anchored to misfolded proteins because the C-terminal diglycine motif is not exposed. A reasonable molecular model to explain the physical interaction between HDAC6 and aggregated proteins has been proposed (Ouyang et al., 2012); the deubiquitinase ataxin-3 is first recruited to aggregated proteins and partially cleaves them through its deubiquitination activity, this then allows HDAC6 to recognize aggregated proteins through those exposed and unanchored ubiquitin C-termini. The contribution of HDAC6 to aggresome formation may be linked to pathogenesis of human diseases; for example, it is concentrated in Lewy bodies, aggresome-related structure observed in neurodegenerative disorders such as Parkinson’s disease.

(3) In addition to dynein, HDAC6 interacts with other proteins and forms a complex to bring aggregated proteins into aggresome. Autophagy receptors including p62/SQSTM1 link aggregated proteins to autophagy. These proteins interact with both ubiquitin and LC3/GABARAP on the autophagosome, thereby promoting autophagy of ubiquitinated aggregated proteins (Kirkin et al., 2009). An E3 ubiquitin ligase, TRIM50 interacts with an autophagy receptor p62/SQSTM1 and also with HDAC6 to promote the sequestration of aggregated proteins into the aggresome (Fusco et al.,

2012). As mentioned above (in **1.2.2**), HDAC6 deacetylates TRIM50 at K372, resulting in its ubiquitination and localization on aggresome (Fusco et al., 2014). It is important to note that the relationship between HDAC6/dynein complex and HDAC6/p62/TRIM50 complex for bringing aggregated proteins into aggresome is still unknown.

(4) HDAC6 does not only contribute to aggresome formation, but also plays important roles for its clearance process. Autophagy had been predominantly defined as a non-selective degradation caused by nutrient starvation. However, it has now become clear that autophagy is also dedicated to cellular quality control by selective disposal of aggregated proteins, aggresomes and damaged organelles (Mizushima and Komatsu, 2011). HDAC6 regulates this quality-control autophagy by actin filament remodeling through deacetylation of cortactin (Lee et al., 2010) (discussed in **1.2.2**). Cortactin is recruited to protein aggregates where it enhances F-actin polymerization to clear aggregates by autophagy. Deacetylation by HDAC6 is required for cortactin to stimulate F-actin polymerization in the process. Thus, HDAC6 KO MEFs are defective in aggregated protein clearance, which is due to inefficient autophagosome-lysosome fusion process, but are normally equipped with starvation-induced autophagy.

(5) Autophagy acts as a compensatory degradation system through aggresome clearance in case that UPS is inhibited, as discussed so far. However, proteasomes are also localized in aggresomes and Lewy bodies although they cannot process protein aggregates. Poh1, a proteasomal deubiquitinase produces C-terminal free unanchored ubiquitin chains in aggresome, resulting in HDAC6 activation there and subsequent quality-control autophagy of aggresome (Hao et al., 2013). Poh1 specifically cleaves K63 linkages (Cooper et al., 2009) and K63-linked, but not K48-linked, ubiquitin chains interact with HDAC6. The K63-linked ubiquitin chains produced by Poh1 bind to HDAC6 ZnF-UBP and activate its catalytic activity, leading to deacetylation of the substrate cortactin. Cortactin-mediated autophagy process does not depend only on F-actin, but also depends on MHC10, as knockdown of MHC10 impairs autophagy of aggresome.

This regulation of aggresome-autophagy pathway by HDAC6, dynein motor and

actinomyosin system is “hijacked” and utilized by influenza virus A (IAV) for its uncoating. This identified HDAC6 as a host factor important for IAV infection (Banerjee et al., 2014). Our group together with the lab of Ari Helenius showed that IAV capsid uncoating process is severely impaired in HDAC6 deficient cells. Analysis of the stepwise IAV entry in HDAC6 deficient cells revealed no significant differences in other processes such as endocytic uptake, acid-induced HA conversion, and fusion. Impairment of capsid uncoating in the absence of HDAC6 is rescued by catalytic-dead mutant of HDAC6, but not by non-ubiquitin-binding mutant, suggesting that ZnF-UBP function is essential for this process. Virions were found to be loaded with C-terminal free ubiquitin chains inside, thus potentially explaining the requirement for the ZnF domain (Banerjee et al., 2014). Inhibition of microtubule/actin filament assembly and dynein/myosin motor system, which are important for aggresome formation by HDAC6, also impairs capsid uncoating. These observations allowed to conclude that HDAC6 works as a linker molecule between ubiquitin chains in IAV and motor proteins, which allows to generate physical forces to break apart the capsid. It still remains unknown which other types of viruses besides IAV have polyubiquitin chains inside and how this process is regulated.

(6) As discussed so far, the action of HDAC6 is required for normal resolution of the protein aggregates. Our previous work (Kwon et al., 2007) highlighted HDAC6 as an important factor for SG formation. HDAC6 was found to bind to a SG component, G3BP and to localize to SGs under several stress conditions, including oxidative stress and UV (Kwon et al., 2007). Moreover, HDAC6 KO MEFs exhibit impaired SG assembly and this phenotype is rescued by re-expression of WT HDAC6, but not by catalytic-dead or non-ubiquitin-binding mutants of HDAC6. SG assembly is impaired by microtubule-disrupting drugs, suggesting that the assembly process is dependent on the microtubule network. These observations encouraged us to propose a model that HDAC6 brings translationally stalled mRNAs binding with G3BP to SGs using dynein motor system on microtubule network, similar to the contribution of HDAC6 to aggresome formation. It is important to mention, however, that there has been no rationale to explain the importance of HDAC6 catalytic activity for SG assembly.

In sum, HDAC6 collects ubiquitinated aggregated proteins by bringing them to aggresome using motor proteins, where the interaction between HDAC6 and ubiquitinated proteins are tuned by chaperones. Ubiquitin chains derived from aggresome processing enhance catalytic activity of HDAC6 to promote re-constitution of actinomysin network surrounding aggresome and supports subsequent quality-control autophagy. The formation of another type of protein condensates, SG is also regulated by catalytic activity of HDAC6 although the molecular mechanism is not clarified yet.

Thus, HDAC6 was known to be important for the formation of different kinds of cellular granules, such as aggresomes or SGs, in response to environmental cues including stress but there was no unifying mechanistic rationale. This therefore remained a long-standing unanswered question until recently, – when the general concept from physics “phase transitions” was introduced into molecular biology by biophysicists (Brangwynne et al., 2009).

1.3 Phase transitions and cellular granules

1.3.1 Liquid-liquid phase separation in biology

Cellular granules (also known as membrane-less organelles or biomolecular condensates) are involved in diverse biological processes, including protein homeostasis, RNA metabolism, ribosome biogenesis, and signal transduction (Fig. 9).

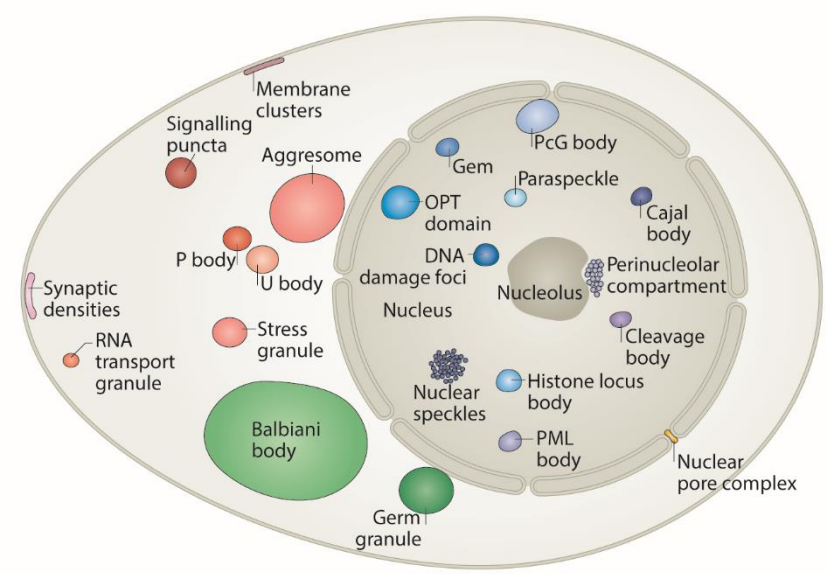


Figure 9. Cellular granules. Schematic of the numerous condensates in the nucleus, cytoplasm and membranes of eukaryotic cells. Some compartments are observed only in specific cell types. For example, Balbiani bodies and germ granules are specific to germ cells (green hues), and RNA transport granules and synaptic densities are specific to neuronal cell types (pink hues). Adapted from (Banani et al., 2017).

Classic cellular organelles, such as mitochondria and the endoplasmic reticulum, are defined as compartments surrounded by lipid membranes. Those membranes physically separate the interior and exterior of the organelles, and transport machineries tightly regulate organelle compositions. In contrast, granules, including aggresomes and stress granules (SGs) in the cytoplasm, as well as nucleoli, and promyelocytic leukemia (PML) bodies in the nucleus, lack a physical barrier to separate their internal components from the surrounding milieu. Fundamental questions – how they form and how

they maintain their structures and compositions – remained elusive for many years due to their complex composition, consisting of dozens of different molecules and the lack of genetics.

The recent discovery, that germ (P) granules in *C. elegans* are liquid-like, brought an important clue to understand the process of granule formation, and it is now recognized that phase transition, especially liquid-liquid phase separation (LLPS, also called coacervation) driven by multivalent macromolecular interactions is contributing critically to the formation of these organelles (Banani et al., 2017; Brangwynne et al., 2015; Hyman et al., 2014; Shin and Brangwynne, 2017). Phase transitions are commonly used to describe transitions between solid, liquid and gaseous states, with each phase having unique physical characteristics (Weber and Brangwynne, 2012). For example, gaseous molecules in water vapor move freely and rarely interact with each other. In liquids, water molecules transiently form hydrogen bonds for interaction, but constantly shuffle their neighborhood. The water molecules in solid phase are ordered in a lattice with stable hydrogen bonds holding neighboring molecules. Macromolecules such as proteins and RNAs *in vivo* also undergo phase transitions. The clearest example is cytoskeletal proteins; actin and tubulin can change their states between a soluble gas-like state where monomers freely diffuse in cytosol and a solid state as the component of filaments. Recent careful biophysical observations of P granules in *C. elegans* gave convincing demonstration that a liquid-like state is appropriate to describe the dynamics of the granule, and maybe those of other cellular granules (Brangwynne et al., 2009). This idea was supported by the following observations (Fig. 10):

(1) P granules deform in shear flows.

(2) The photobleached region of a P granule will recover through internal rearrangement.

(3) Two spherical P granules fuse together, and form a new spherical shape.

Thus, P granule formation appeared to be governed by LLPS, which is often seen in the daily life; for example, water and oil (two immiscible liquids) always phase separate. A decade of studies subsequently have put emphasis on the general importance of LLPS for formation of biomolecular condensates where components for biochemical reactions such as protein and RNA spatiotemporally

localize to achieve a certain biological output. This concept is well exemplified in the formation of various types of cell signaling hubs, including centrosome and heterochromatin (Banani et al., 2017).

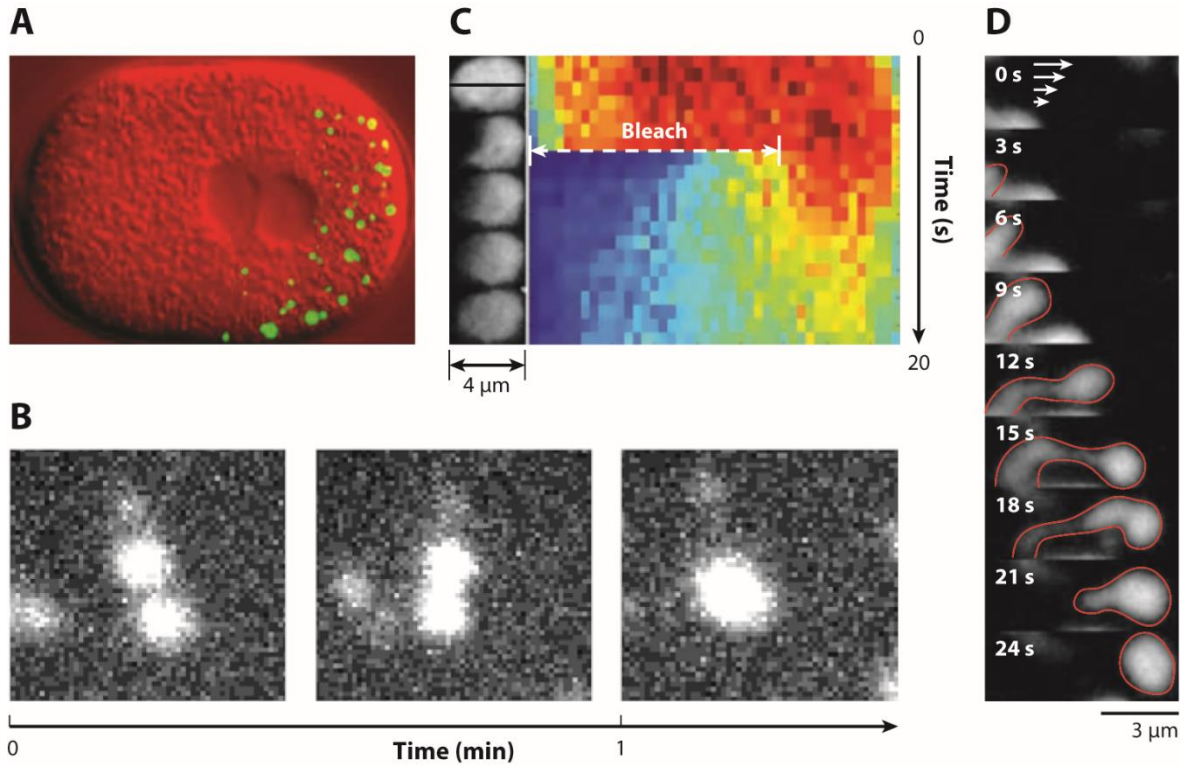


Figure 10. P granules exhibit liquid-like properties. (A) P granules (green) in the cytoplasm of a one cell-stage *C. elegans* embryo. (B) Two P granules fuse and relax their shape within about one minute. (C) FRAP of P granule (left). Kymograph of linear intensity profiles along the anterior-posterior axes (right). Red color indicates high intensity and blue corresponds to background intensity. (D) P granule (red outline) deformed by sheared flow with a direction indicated by the white arrows. Modified figure based on (Hyman et al., 2014).

The contribution of LLPS to biological processes in living cells is only emerging. Beyond classical membrane-less organelles, biological condensates are increasingly observed in diverse cellular processes: signaling to the cytoskeleton (Li et al., 2012), T cell activation (Su et al., 2016), neurodegenerative disease (Molliex et al., 2015; Patel et al., 2015), innate immune signaling (Du and Chen, 2018), synaptic vesicle transport and plasticity (Milovanovic et al., 2018; Zeng et al., 2018; Zeng et al., 2016), cell division (Woodruff et al., 2017), heterochromatin formation (Larson et al.,

2017; Strom et al., 2017), transcription (Cho et al., 2018; Chong et al., 2018; Lu et al., 2018; Sabari et al., 2018), RNA splicing (Gueroussov et al., 2017; Ying et al., 2017), miRNA-induced silencing (Sheu-Gruttadauria and MacRae, 2018), RNAi inheritance (Wan et al., 2018), nutrient response (Delarue et al., 2018), autophagy (Sun et al., 2018) and even photosynthesis (Freeman Rosenzweig et al., 2017). At phase transitions, a system tends to go to the lowest energy state, which is determined by entropic and enthalpic factors. Entropy generally tends to keep a system at well mixed state, and thus, the system is usually required to counteract this by some attractive interactions between molecules for phase separation. Studies about phase transitions in nonliving materials have suggested that polymers are prone to undergo phase separation because of connectivity of individual subunits. This implicates the importance of biopolymers, such as protein, RNA, and DNA in intracellular phase separation. It is now widely acknowledged that distinct proteins and RNAs with certain propensities contribute to the process. The three representative molecular determinants – multivalent proteins with modular interaction domains, intrinsically disordered proteins (Fig. 11), and repeat-containing RNAs – are well known; they usually form phase-separated liquid droplets at certain external conditions (e.g. temperature, salt concentration and pH) *in vitro*, and this activity is well correlated with membrane-less organelle formation *in vivo*. These biomolecules with phase separation propensity are listed (Table 4) below, and discussed in more detail in the following section.

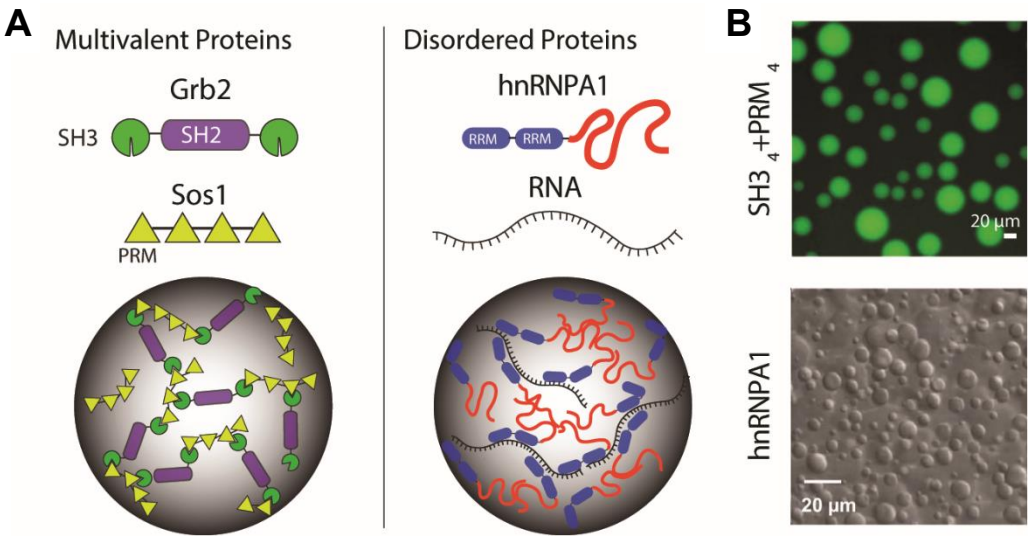


Figure 11. Protein features driving intracellular phase separation. (A) Weak interactions between multivalent proteins drive LLPS. Proteins composed of signaling pathways often contain a set of binding domains, such as SH3 and proline-rich motif (PRM). Disordered proteins are also key for driving LLPS. Many RNA-binding proteins have IDRs in addition to the domain for RNA recognition such as RNA recognition motif (RRM). (B) (Top) A mixture of two synthetic proteins, repeats of SH3 and PRM, undergoes LLPS. (Bottom) The purified hnRNPA1, an IDR-containing RNA-binding protein, also forms liquid droplets. Adapted from (Shin and Brangwynne, 2017).

Proteins/RNAs tested for its phase separation propensity	Molecular determinants	Relevant biomolecular condensates/ biological outcomes	Reference
1. Nuclear pore proteins (e.g. Nsp1)	FG-rich IDR	Nuclear pore complex	(Frey et al., 2006; Labokha et al., 2013)
2. N-WASP signaling components	SH2 domain + pY, SH3 domain + PRM	Remodeling of actin cytoskeleton networks	(Li et al., 2012)
3. RPB1 (subunit of RNA pol II)	Heptapeptide (YSPTSPS) repeat in its C-terminal domain	Transcription	(Kwon et al., 2013)
4. C9orf72 (pathogenic mutants)	Dipeptide repeat (GRn, PRn)	Stress granules, nuclear speckles and Cajal bodies ALS/FTD	(Boeynaems et al., 2017; Kwon et al., 2014; Lee et al., 2016)
5. DDX4	FG, RG-rich IDR	P granules	(Nott et al., 2015)
6. LAF-1	RG-rich IDR	P granules	(Elbaum-Garfinkle et al., 2015)
7. FUS (and its family proteins)	G/S-Y-rich (Prion-like) LCS, RGG/RG domain, and RRM	Stress granules ALS/FTD	(Murakami et al., 2015; Patel et al., 2015)
8. hnRNPA1/A2	G/S-Y, RG-rich (Prion-like) LCS and RRM	Stress granules ALS/FTD and MSP	(Molliex et al., 2015; Xiang et al., 2015)
9. RNA binding proteins (Pub1, Lsm4 and eIF4GII)	IDR and RNA binding	Stress granules and P bodies	(Lin et al., 2015)
10. Whi3	QN-rich IDR and RNA binding to its RRM	Spatial patterning of transcripts in the fungus *Ashbya gossypii*	(Langdon et al., 2018; Zhang et al., 2015a)
11. LAT signaling components	SH2 domain + pY, SH3 domain + PRM	T cell receptor signaling pathways	(Su et al., 2016)
12. SPOP + DAXX	Oligomerization domain (SPOP) + SPOP-binding motifs in IDR (DAXX)	Nuclear speckles, PML bodies and DNA-damage loci Prostate cancer	(Bouchard et al., 2018; Marzahn et al., 2016)

13. NPM1	RNA binding to its RRM	Nucleoli	(Feric et al., 2016)
14. Xvelo	Prion-like domain	Balbiani body	(Boke et al., 2016)
15. PML body components	SIM + SUMO	PML bodies	(Banani et al., 2016)
16. PSD95	Coiled-coil domain	Postsynaptic densitiy	(Zeng et al., 2016)
17. TDP-43	α-helix	Stress granules, cellular aggregates ALS/FTD	(Conicella et al., 2016)
18. PGL-1 and PGL-3	RGG/RG domain	P granules	(Saha et al., 2016; Zhang et al., 2018a)
19. Dhh1	RNA binding	P bodies	(Mugler et al., 2016)
20. Pab1	RRM and P-rich LCS	Stress granules	(Riback et al., 2017)
21. SPD-5	Coiled-coil domain	Centrosome	(Woodruff et al., 2017)
22. Repeat-containing RNAs (Triplet repeat: CAG, CUG and G_4C_2 repeat)	RNA-RNA base-pairing	Nuclear speckles Neurodegenerative diseases (Huntington's disease, ALS/FTD and SCA) and myotonic dystrophy	(Jain and Vale, 2017)
23. HP1α	IDR	Heterochromatin	(Larson et al., 2017; Strom et al., 2017)
24. Rbfox	G/S-Y-rich LCS	RNA splicing	(Ying et al., 2017)
25. TIA1	Q-rich (Prion-like) LCS	Stress granules ALS/FTD and myopathy	(Mackenzie et al., 2017)
26. Tau	IDR	Neurodegenerative diseases (Alzheimer's disease and Parkinson's disease)	(Hernandez-Vega et al., 2017)
27. Sup35	Prion domain	pH sensor to drive phase separation of translation termination factor in yeast	(Franzmann et al., 2018)
28. UBQLN2	IDR and ubiquitin binding to its UBA	Stress granules	(Dao et al., 2018)
29. Ago2 + TNRC6B	PIWI domain + GW-rich IDR	Assembly of miRNA-induced silencing complex	(Sheu-Gruttadauria and MacRae, 2018)
30. p62/SQSTM1	Ubiquitin binding to its UBA	Autophagy	(Sun et al., 2018)
31. Htt	PolyQ tract	Cellular aggregates	(Peskett et al., 2018)

(pathogenic mutants)		Huntington's disease	
32. CYCT1	H-rich IDR	Transcription	(Lu et al., 2018)
33. DYRK1A	H-rich IDR	Transcription	(Lu et al., 2018)
34. BRD4	IDR	Transcription, component of super-enhancers	(Sabari et al., 2018)
35. MED1	S-rich IDR	Transcription, component of super-enhancers	(Sabari et al., 2018)
36. cGAS	DNA binding	Innate immunity	(Du and Chen, 2018)
37. Synapsin1	IDR with PRM	Presynaptic site	(Milovanovic et al., 2018)
00. DDX3X	RG-rich IDR	Stress granules	(Saito et al., 2019)

Table 4. Summary of proteins having phase separation propensity. Proteins/RNAs whose phase separation propensity were confirmed by *in vitro* droplet formation and/or hydrogel formation are listed in chronological order. It is important to note that there is no consensus on the definition of intrinsically disordered region, low complexity sequence and prion-like domain, because this is dependent on the prediction algorithms. It is important to note that LLPS of 00. DDX3X is confirmed by the work for this thesis, and it is separately discussed in detail (**2.1**). IDR, intrinsically disordered region; PRM, proline rich motif; LCS, low complexity sequence; ALS/FTD, amyotrophic lateral sclerosis/frontotemporal dementia; RRM, RNA recognition motif; MSP, multisystem proteinopathy; SIM, SUMO-interacting motif; SCA, spinocerebellar ataxia UBA, ubiquitin-associated domain.

1.3.2 Molecular determinants in cellular phase transitions

Modular interaction domains for multivalency

Multivalent protein-protein interactions by modular interaction domains are the first key molecular determinant for phase transitions. Actin-regulatory signaling pathway consisting of three proteins: nephrin, Nck and neural Wiskott-Aldrich syndrome protein (N-WASP) demonstrated the contribution of LLPS driven multivalent proteins to cell signaling (Li et al., 2012). These proteins form oligomers through two types of interactions: Src homology 2 (SH2) domains in Nck and phospho-tyrosine residues (ligands of SH2) in nephrin, and SH3 domains in Nck and proline-rich motifs (PRMs, ligands of SH3) in N-WASP. The resulting assembly shows both *in vitro* phase-separated liquid droplets and clusters on lipid bilayers in cells. A similar mechanism is observed in T cell receptor signaling pathways (Su et al., 2016). In this case, five proteins: linker for activation of T cells (LAT), growth factor receptor-bound protein 2 (GRB2), GRB2-related adaptor protein 2 (GADS), son of sevenless (SOS) and SH2 domain-containing leukocyte protein of 76 kDa (SLP76), which form multivalent interactions with each other through their SH2, SH3 domains and ligands, assemble into membrane puncta in cells in response to stimulation of the T cell receptor.

Various synthetic proteins with repeated domains have been utilized to delineate the fundamental principles governing intracellular LLPS. Due to the complex nature of the diversity of proteins within granules, it is still challenging to elucidate the key molecular parameters during granule formation in cells. Simplified model systems composed of a few types of proteins, which harbor a variable number of modules (binding domain and its ligand), have been developed so far to establish generalizable concepts. One example is the system with repeats of a single SH3 domain ((SH3)m, m: number of SH3), and repeats of proline-rich motif ligand ((PRM)n, n: number of PRM) (Li et al., 2012). These two protein species have identified valency (that is, the number of interaction modules) and affinity between two modules. In this system, if repeats of four or more SH3 domains ($n \geq 4$) are mixed with similarly repeating PRM domains, liquid droplets form. Moreover, the concentration of protein species at which droplets are observed decreases with higher valency; this clearly shows that the

concentrations needed for LLPS are directly related to the valency of each protein species. Another system with small ubiquitin-like modifier (SUMO) ligand and the SUMO interacting motif (SIM) provided important insights in the principles of compositional control of granules (Banani et al., 2016). As observed in the SH3/PRM system, when higher valency species ((SUMO)n and (SIM)m) are mixed, they undergo LLPS, forming liquid droplets. If there are more unbound free SUMO ligands within droplets (when n > m), SIM species are preferentially recruited. On the other hand, when n < m, SUMO ligands are strongly recruited. Taking into account that SUMO-SIM interactions appear to be critical for the recruitment of many PML body components, this simple model system may help to understand the complexity of actual cellular granules.

Molecular assembly and phase separation of multivalent proteins are distinct phenomena, but some proteins with coiled-coil domain undergo LLPS through their oligomerization. Postsynaptic densities (PSDs) are protein enriched compartments under postsynaptic membranes, whose formation and activity-dependent modulation is considered as the most basic mechanisms regulating synaptic plasticity in the nervous system. The PSD protein SynGAP, a Ras GTPase activating protein binds to another postsynaptic density protein, PSD95, resulting in LLPS of the complex, forming liquid-like droplets reminiscent of the PSDs. (Zeng et al., 2016). There, SynGAP forms a parallel coiled-coil oligomer and binds to PSD95 molecules in a 3:2 stoichiometry. This LLPS is important for SynGAP stabilization in PSDs and for preventing neurons from hyperexcitation. This finding provides important insight that oligomerization of one protein species further leads to interactions with other proteins for LLPS, and explains its role in formation of PSDs. Centrosome formation appears to be triggered by LLPS of another coiled-coil protein, SPD-5 (Woodruff et al., 2017). Centrosomes, membrane-less organelles for nucleating microtubule arrays, consist of centrioles surrounded by a dynamic assembly of protein called the pericentriolar material (PCM). SPD-5 harbors nine predicted coiled-coil domains, which make up ~40% of the protein. It forms spherical condensates *in vitro* with other PCM components, the microtubule polymerase ZYG-9 and the microtubule-stabilizing protein TPXL-1. This condensate recruits tubulin, and reconstitutes nucleation of microtubule asters; it exactly

resembles PCM *in vivo*. Remarkably, ZYG-9 itself is sufficient for microtubule nucleation in the system, but the formation of asters requires three proteins together in condensates, indicating the importance of LLPS in centrosome formation.

Intrinsically disordered proteins

Intrinsically disordered regions (IDRs) and proteins that harbor these regions (intrinsically disordered proteins, IDPs) represent the second key molecular determinant that can phase separate under physiological conditions. IDRs do not have a defined 3D structure but often contain sequence elements for multivalent weak intermolecular interactions. IDPs are enriched in various granules, especially in ribonucleoprotein (RNP) granules, including P granules, P bodies, and SGs. Many such proteins can phase separate on their own *in vitro* under various solution conditions. The sequences of IDRs generally lack hydrophobic residues for folding, and they are typically enriched in particular amino acids, including polar (G, Q, N, S), positively charged (R, K), negatively charged (D, E), or aromatic (F, Y). IDRs can be divided into substructures based on their size and function (Calabretta and Richard, 2015); short linear motifs (SLiMs, 1-10 amino acids), which act as ligands for protein domains or serve as consensus sequences for enzymes, molecular recognition features (MoRFs, 10-70 amino acids), which become structured after protein binding, and low complexity sequences (LCSs, hundreds of amino acids), which undergo phase transitions under physiological condition. The lack of sequence diversity in LCS yields multiple repetitive motifs such as -G/S-Y-, -QN-, -FG-, -RGG- in these molecules, which enable polypeptide backbone to interplay through intramolecular/intermolecular electrostatic interactions (Fig. 12). Neurodegenerative diseases (e.g. ALS/FTD)-related RNA binding proteins often have pathogenic mutations in their IDRs. FUS, hnRNPA1 and hnRNPA2, for example, harbor LCSs and their phase transitions have been intensively studied so far (Kato et al., 2012; Molliex et al., 2015; Murakami et al., 2015; Xiang et al., 2015).

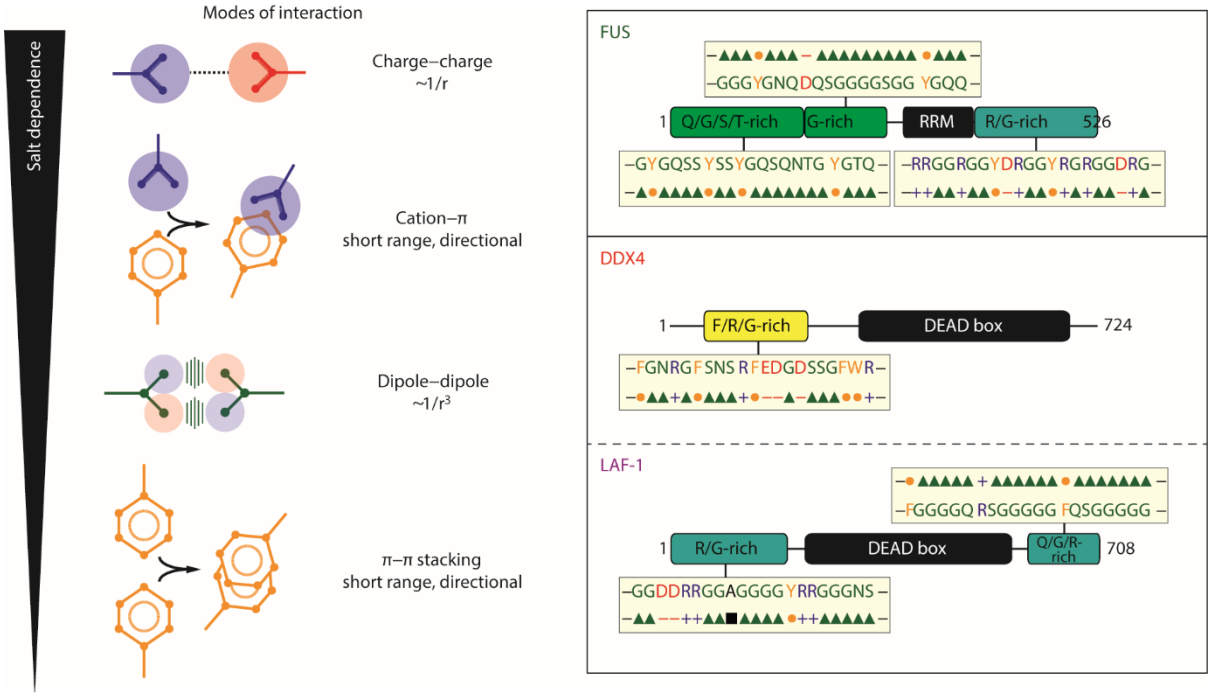


Figure 12. Molecular interactions underlying phase transitions. The three panels at the right show the domain architectures for FUS, DDX4 and LAF-1. The amino acid biases within LCSs of each protein are highlighted in square boxes that show the partial amino acid sequence and annotation of the sequence stretch using symbols to denote dipoles (triangles), charges (+ or -), aromatic groups (circles) and aliphatic groups (squares). The panel at the left shows schematics of the hierarchical modes of interactions that may contribute to phase separation. This is annotated to depict the types of interactions, their ranges, and their expected salt dependencies. Adapted from (Brangwynne et al., 2015).

Their LCSs are commonly characterized by repetitive -G/S-Y- motifs, and also often recognized as a type of prion-like domains (PrLDs): a class of LCS particularly enriched with uncharged polar amino acids (Y, N, and Q) and also resembling the amino acid frequency in yeast prion domains (Aguzzi and Altmeyer, 2016). PrLDs are prone to aggregation and form amyloid-like fibers, and this is experimentally well-confirmed. Phase-separated droplets of FUS, hnRNPA1 and hnRNPA2 proteins rapidly exchange molecules with the soluble phase, but these proteins can form hydrogels comprised of amyloid-like fibers. Several recent studies have commonly revealed that they form liquid droplets at moderate concentrations, but these droplets can nucleate amyloid-like fibers and become hydrogel

over time under high concentrations and low-temperature conditions. This process is accelerated when these proteins harbor disease-associated mutations. This liquid-"solid" phase transition in context of physiology and pathology will be discussed later (discussed in **3.3.1**).

ALS/FTD caused by C9orf72 is the most drastic example to demonstrate that pathogenic mutations can alter the propensity of LLPS and granule formation. Expansion of a hexanucleotide repeat GGGGCC (G_4C_2) in C9orf72 is the most common cause of C9orf72-ALS/FTD (C9-ALS/FTD). Unaffected individuals typically harbor between 2 and 23 G_4C_2 repeats, while people with C9-ALS/FTD harbor hundreds to thousands of repeats. This mutation appears to affect disease progression through two types of toxic gain of function mechanism; one is from abnormality of the transcripts itself, and the other is from dipeptide repeat (DPR) proteins produced by repeat-associated non-ATG translation, which occurs without an initiating AUG codon and produces peptides from all reading frames. Five DPR proteins are produced from pathogenic C9orf72: GAn, GRn, PRn, PAn and GPn (n: number of DPR, $n \geq 20$ are usually utilized as representative protein species). All of these DPRs are produced in C9-ALS/FTD patients and account for cytoplasmic and nuclear inclusions observed in the brain and spinal cord. Some DPR species, those containing arginine (GRn and PRn) are especially toxic, and the role of DPRs in LLPS and granule formation has recently become clear. PRn itself form liquid droplets and it also enhances LLPS of other IDPs (Boeynaems et al., 2017). Remarkably, GRn or PRn expression affects intracellular phase separation; they bind to LCSs of the proteins, and impair the dynamics of several types of granules, including SGs, nucleoli, nuclear speckles and Cajal bodies (Kwon et al., 2014; Lee et al., 2016; Lin et al., 2016).

Apart from pathogenesis of neurodegenerative diseases, LCSs comprised of repetitive -G/S-Y- motifs also contribute to a fundamental process for mRNA maturation: alternative splicing. Alternative protein isoforms and their structures from the database suggest that alternatively spliced exons, including those subject to tissue-, species- and lineage- specific regulation, are significantly enriched in IDRs. It is therefore assumed that the evolution of alternative splicing of IDRs has affected biological complexity by remodeling protein interaction through IDRs. Studies about evolution of

alternative exons within vertebrates have revealed that important insights for regulation of hnRNPs and other RNA binding proteins (RBPs) (Gueroussov et al., 2017); mammalian-specific exons are concentrated in the -G/S-Y- rich LCSs of hnRNPs and other RBPs. Differential inclusion of the alternative exons, which affects the number of -G/S-Y- motifs in LCSs, regulates formation of hnRNP complexes through changes in their LLPS propensity. Formation of these hnRNP complexes on pre-mRNA, in turn, is critical for the global regulation of alternative splicing events. This regulatory mechanism is also observed in other RBPs, Rbfox proteins; higher-order assembly of Rbfox with large assembly of splicing regulators (LASR) proteins is required for their large-scale exon activation during splicing, and this interaction is mediated by LLPS through the Rbfox -G/S-Y- rich LCS (Ying et al., 2017).

As exemplified above, IDPs related to ribostasis often undergo LLPS. ATP-dependent DEAD-box RNA helicase families, for example, harbor their conserved central helicase domain, and IDRs at their N- and C-terminus. One of the family members, DDX4 is a major component of nuage/chromatoid body, which hosts machineries of an RNAi pathway, protecting spermatocytes and spermatids from transposable elements. The N-terminal IDR of DDX4 alone is sufficient for LLPS, and $DDX4^{YFP}$, where the helicase domain is substituted by YFP, forms droplets within the cell nucleus (Nott et al., 2015). Another family members, LAF-1 (*C. elegans* homolog of DDX3) forms similar droplets *in vitro* and contributes to P granule assembly (Elbaum-Garfinkle et al., 2015). The IDRs within DDX4 and LAF-1 are polyampholytic; they contain roughly equivalent fractions of positively and negatively charged amino acids (Das et al., 2015). Although, the mechanism underlying charge pattern-driven LLPS is still poorly understood, -FG- and -RGG- motifs appear to be important in case of DDX4 and LAF-1, respectively.

Recent studies demonstrated that heterochromatin formation is also regulated through LLPS (Larson et al., 2017; Strom et al., 2017). A heterochromatin protein 1 (HP1) family proteins, HP1α, can form liquid droplets, which is unique to this protein among HP1 family members. It appears that two types of IDRs of HP1α, hinge region between chromo and chromoshadow domains and HP1α-

specific N-terminal extension (NTE), contribute to LLPS. This LLPS is induced only by phosphorylation of NTE or presence of DNA, which allow the protein to change its conformation and bind other HP1 proteins. Heterochromatin domains, indeed, exhibit dynamics that are characteristic of LLPS; HP1 initially diffuse, then forms spherical foci that grow, fuse together, and dissolve when it is removed from chromatin. These findings suggest that heterochromatin-mediated gene silencing may occur through sequestration of compacted chromatin in HP1 liquid droplets. In addition to heterochromatin formation, LLPS may be related to transcriptional hub formation; thus, it is emerging as a general mechanism for transcriptional regulation (discussed in **3.3.2**).

Some proteins have both modular interaction domain and IDR to promote phase separation cooperatively. LLPS also plays a role in synaptic vesicle (SV) clusters in proximity to presynaptic sites, as well as PSDs mentioned above. These clusters act as a reservoir from which SVs are drawn for exocytosis during prolonged synaptic activity. Synapsin1, a component of the matrix for crosslinking SVs, interacts with SH3-containing proteins through its IDR harboring PRMs, and undergoes LLPS. In this case, the IDR itself is sufficient for LLPS, but it is enhanced by co-existence of SH3 domains. The synapsin1 droplets could capture small lipid vesicles (50-150 nm diameter), and they are dispersed upon phosphorylation by calcium/calmodulin-dependent protein kinase II (CaMKII) (Milovanovic et al., 2018). Thus, these *in vitro* re-constitution experiments mimic well the release of SVs occurring at presynaptic sites upon stimulation, suggesting that LLPS plays a role in the process (Milovanovic and De Camilli, 2017). This concept – vesicle clustering into a phase-separated liquid phase – may apply to other membrane-bound vesicles involved in traffic among ER, Golgi and endosome, and will help to understand their dynamic organization in future.

Multivalent protein-protein interactions mentioned above are often modulated by posttranslational modifications (PTMs) to alter the phase separation threshold, as observed in the case of synapsin1. PTMs can either change the confirmation, directly generate binding motifs (e.g. tyrosine phosphorylation as a ligand of SH2 domains), or increase or decrease the net charge of the IDRs above or below the threshold. So far, it has been reported that methylation, phosphorylation, sumoylation and

non-covalent interactions between polyubiquitin/poly(ADP-ribose) and their target proteins regulate phase separation (discussed in **3.2.4**).

RNAs as a scaffold for multivalent interactions

RNAs form the third key molecular determinants contributing to phase separation under physiological conditions, whose major role appear to promote phase separation of multivalent RNA-binding proteins and IDPs. Polypyrimidine tract binding protein (PTB), for example, is a multivalent RNA-binding protein with four RNA recognition motif (RRM) domains (Li et al., 2012). The nucleolar protein nucleophosmin (NPM1) has a RRM domain, and requires binding with ribosomal RNA (rRNA) to form multivalent interactions that drive phase separation (Feric et al., 2016). These proteins undergo LLPS in the presence of RNA, independent of any IDRs. In parallel, however, multivalent and weak interactions between IDRs and RNAs additionally contribute to LLPS. The LLPS of N-terminal IDR of LAF-1, which is enriched with RGG- motifs, is affected by RNA; RNA increases fluidity and dynamics within LAF-1 droplets (Elbaum-Garfinkle et al., 2015). In many cases, RNA-binding proteins contain both RNA-binding motifs and IDRs, and thus it is assumed that interactions with RNA promote LLPS synergistically with IDR-IDR interactions. This is well described by the observation that the saturation concentration for LLPS of chimeric PTB proteins, where PTB is fused with each IDR from different RNA-binding proteins (Pub1, Lsm4, eIF4GII, Tia1, hnRNPA1 and FUS), is significantly lowered in the presence of RNA (Lin et al., 2015).

The different chemical/structural features which are yielded from genetic code information of mRNAs impact phase transitions. A pioneer work with Whi3 protein, which accommodates an IDR with polyQ and a RRM domain, indicates that different RNAs may induce a variety of granules (Zhang et al., 2015a). Whi3 assembly is required for organizing both CLN3 mRNA at sites of nuclear division and BNI1 mRNA at polarity sites of the filamentous fungus *Ashbya gossypii*. Interestingly, these two different mRNAs individually let Whi3 phase separate with different thresholds, and form droplets with different physical properties. The subsequent work clarified that secondary structures of CLN3

and BNI1 mRNAs define the valency of Whi3 binding to these mRNAs, resulting in discrete properties of two droplets; perinuclear droplets are less dynamic with low protein/RNA ratio while polarity droplets are more dynamic with high protein/RNA ratio (Langdon et al., 2018).

PolyQ Whi3 droplet with two types of mRNAs clearly demonstrates that the interaction between protein and RNA, and more specifically that the ratio of protein to RNA in the complex governs phase separation *in vivo*. This principle may be applied to other RNA binding proteins more generally; high protein/RNA ratios promote phase separation, whereas low ratios prevent it (Maharana et al., 2018). According to this rule, the nucleus can be considered as a buffered system where high RNA concentrations keep RNA-binding proteins soluble and prevent their aggregation. Altered concentration of RNA or RNA-binding proteins induces aberrant phase transitions. This concept explains well why disease-related RNA-binding proteins such as TDP-43 and FUS are largely soluble in the nucleus, but form solid pathological aggregates in the cytoplasm. The contribution of RNAs to phase separation discussed so far depends on the interaction with proteins. Recent work, however, proposes the idea that the RNA itself can phase separate and could play a role in neurodegenerative diseases associated with RNA repeats (Jain and Vale, 2017). As mentioned pathogenic G_4C_2 repeat in ALS/FTD above, expansions of short nucleotide repeats cause several neurological and neuromuscular disorders; other examples are nucleotide triplets, including CAG in Huntington disease, and CTG in myotonic dystrophy. These transcripts initially undergo LLPS through their multivalent base-pairing, and rapidly become cross-linked into gels (Fig. 13). They co-localize in nuclear speckles in cells, and can be dissolved by reagents that disrupt RNA gelation, suggesting that the transcripts form granules in the same way as *in vitro* gelation. It is assumed that the aberrant granules sequester various RNA-binding proteins, and cause a disruption of cellular homeostasis. Therefore, these findings may open up new avenues for treating repeat expansion diseases. As discussed further in **1.3.3**, the nature of intracellular phase separation – it is caused by interactions among dozens of proteins and RNAs – has important biological implications, for example in SGs formation.

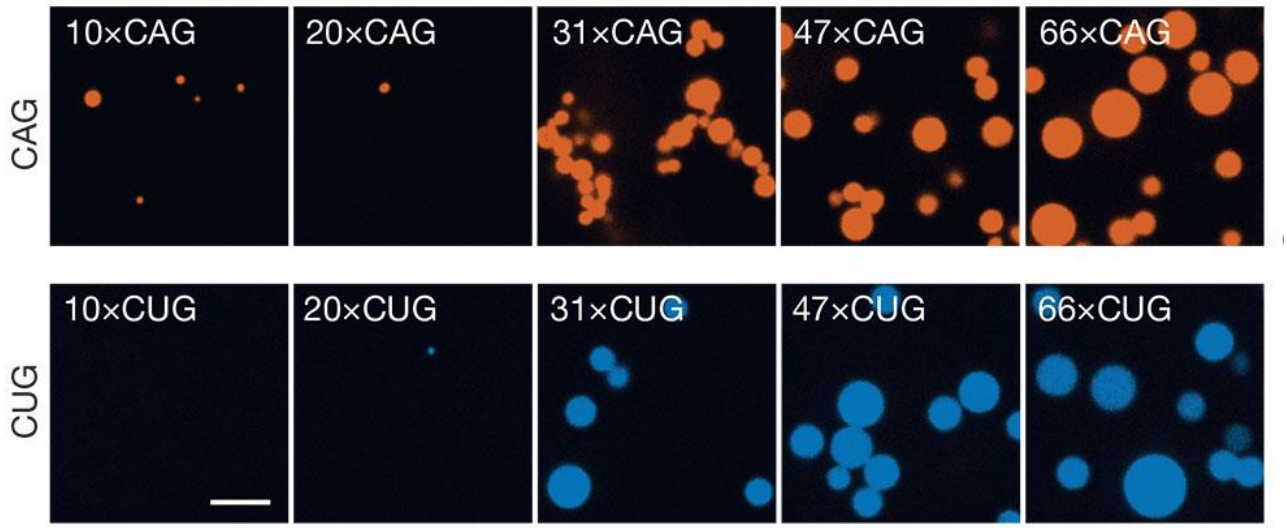


Figure 13. Repeat-containing RNAs form RNA gels. Consistent with the valency dependence, the formation of CAG/CUG RNA clusters occurred only with more than 30 triplet repeats. Adapted from (Jain and Vale, 2017).

Molecular crowding as a general regulator for phase separation

The three molecular determinants mentioned above define the molecular properties that yield weak transient interactions among molecules for phase separation. A crucial determinant for phase separation in another dimension in the system is molecular crowding. For example, it is well known that phase-separated droplet formation *in vitro* is strongly enhanced by adding molecular crowder such as polyethylene glycol (Lin et al., 2015; Molliex et al., 2015). Given that overexpression of some SG component spontaneously formed SGs (Tourriere et al., 2003), increased local concentration of the proteins or RNAs by macromolecular crowding may also play a key role in cellular phase separation. However, the physiological mechanisms that control crowding within the cell and the effects of crowding on phase separation *in vivo* had remained obscure. A recent ingenious work provided the way to measure macromolecular crowding in cells using tracer particles, and obtain insight for cellular phase separation; genetically encoded multimeric nanoparticles (GEMs) were developed to address this question (Delarue et al., 2018). These GEMs are GFP-tagged tracer particles of a defined shape and size (20nm in diameter), which serve as a standard microrheological tool in cells. Using this tool

in combination with *in situ* cryo-electron tomography, the authors found that the mTORC1 pathway controls the diffusion of particles > 20 nm in size nearly 2-fold by regulating ribosome concentration. Moreover, they examined the behavior of the established SUMO/SIM model (Banani et al., 2016) (discussed above), which forms phase-separated droplets at appropriate conditions. Indeed, mTORC1 inhibition by rapamycin attenuated ribosomal crowding and inhibited droplets formation. Thus, ribosomes can act as a crowding agent to induce phase separation, highlighting the general importance of molecular crowder also for cellular phase separation.

1.3.3 Stress granules in physiology and diseases

Various membrane-less organelles have been discovered so far, and some of them are now known as RNP granules due to their composition of RNA and protein, including cytoplasmic granules such as processing bodies (P bodies) and stress granules (SGs) (Fig. 14).

P-bodies are composed of the pool of nontranslating mRNA and the conserved core of proteins involved in mRNA decay and translation repression. They include decapping enzyme complex (e.g. Dcp1/Dcp2), decapping activators (Lsm1-7 complex), translation repressors (Dhh1/RCK/p54, Pat1 and Scd6/RAP55), the 5' to 3' exonuclease (Xrn1), and deadenylase complex (e.g. Ccr4/Pop2/Not). mRNAs within P bodies can be targeted for decapping and degradation by these factors. Translation initiation factors and ribosomal proteins are normally excluded from P bodies. Although they are already present in unstressed cells, P-bodies are induced in response to stresses which inhibit translation initiation.

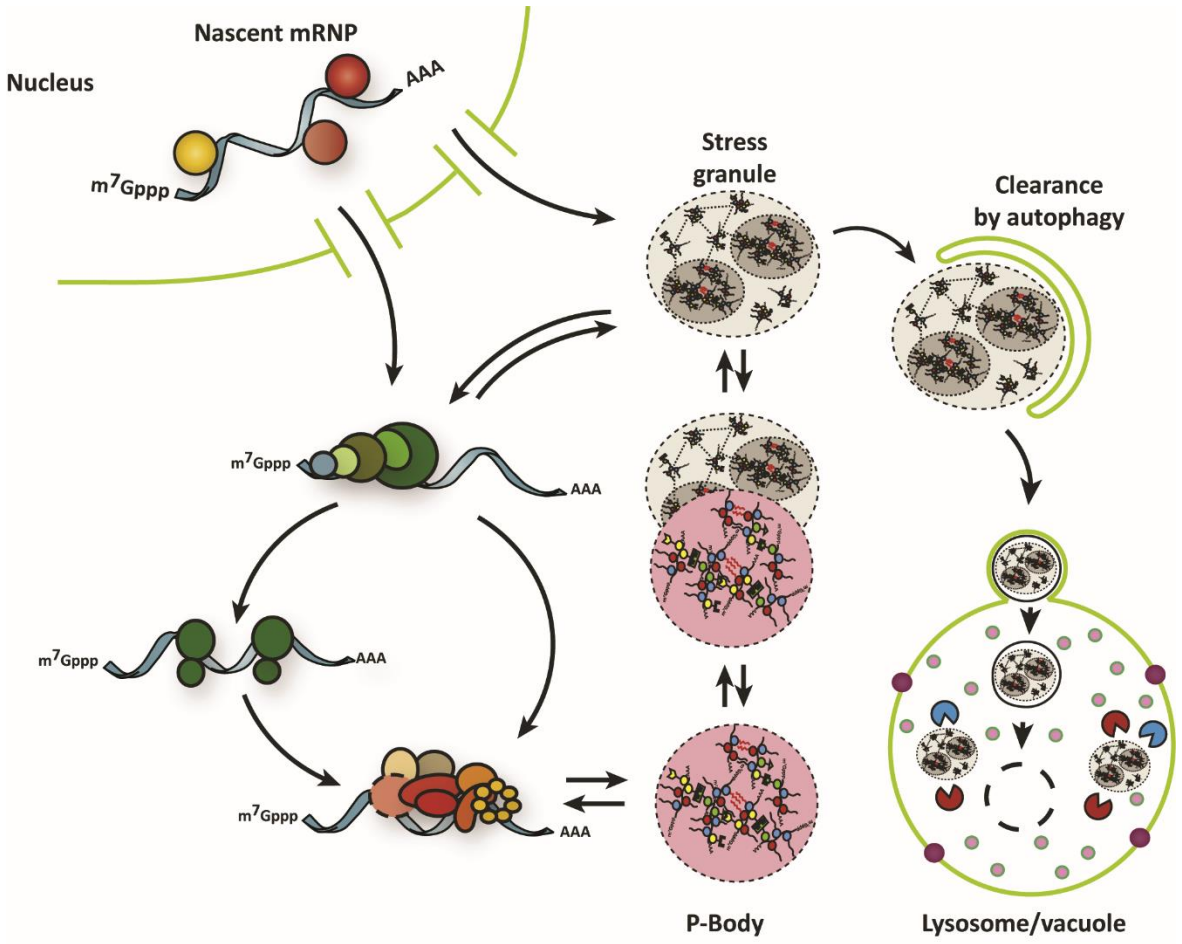


Figure 14. SGs and P-bodies in ribostasis. SGs form from untranslating messenger ribonucleoproteins (mRNPs). SGs can interact with P-bodies, exchange components with the cytoplasm, and undergo autophagy. Adapted from (Protter and Parker, 2016).

SGs are a second type of cytoplasmic RNP granule that can be overlapped with P-bodies. P bodies and SGs can dock to each other, suggesting a dynamic cycle where the components can be transferred and remodeled between these two granules. Similar to P bodies, assembly of SGs is dependent on the pool of nontranslating mRNAs. Although they share some components, SGs especially contain translation initiation factors such as components of eIF2, eIF3 and eIF4 complexes, poly-A binding protein (PABP), and the 40S ribosomal subunit. Assumed from that SGs harbors translation initiation factors, there is a clear link between SGs and the process of translation initiation. SGs indeed form when translation initiation is inhibited either by drugs or by stresses. SGs don't have

a uniform structure, but contain internal substructures identified by super-resolution microscopy. These substructures are referred to as “SG cores” and can be purified biochemically, whereas the remaining portion of the SG is called “SG shell” (Jain et al., 2016). Thus, SGs have two distinct layers: a core structure surrounded by a less concentrated and potentially more dynamic shell. In the prevalent model, SG formation has been considered as a two-step process. The initial step is formation of a dense stable SG core. Some core components harbor IDRs (as mentioned below), and thus, the core could induce accumulation of proteins containing IDRs at high concentration, which then trigger LLPS with other IDR-containing proteins there to form the peripheral shell structure. Proteomic analysis of SG cores revealed that they contain PTM enzymes, metabolic enzymes, chaperone proteins as well as translation initiation factors and RNA-binding proteins, which often contain both RNA recognition motifs and IDRs. A recent study added karyopherins for nuclear transport to this list of SG components (Zhang et al., 2018b). These proteomic surveys also revealed that the composition of SGs can vary under different conditions, such as cell and stress types (Markmiller et al., 2018). In any case, SGs are not mere static aggregates, but rather function as dynamic signaling hubs communicating a ‘state of emergency’ to other signaling pathways, which regulate metabolism and growth to survive in stress conditions.

Similar to other membrane-less organelles, SGs are dynamic structures and show liquid-like properties; they undergo fusion, fission, and flow in the cytosol. Fluorescence Recovery After Photobleaching (FRAP) experiments demonstrate that SG components exchange rapidly, normally with half-time for recovery of less than 30 seconds. At the same time, these FRAP experiments have also revealed an immobile fraction of proteins in SGs, suggesting that a fraction of the molecules within SG components exchange slowly. A fascinating possibility is that a shell fraction exchanges rapidly while a core fraction does slowly. Although most SG components reside in SGs transiently (a few seconds), SGs themselves persist for minutes to hours. Notably, it was recently suggested that there is substantial crosstalk between quality control of the transcriptome (ribostasis) and of the proteome (proteostasis), which are related to SG and aggresome formation, respectively. For example,

various ATP-dependent remodeling complexes (e.g. chaperones) support both protein folding and also RNP granule disassembly. As mentioned above (in **1.2.3**), both aggresome and SGs are cleared by autophagy from cells. This crosstalk may explain the similar clinical outcome in the spectrum of degenerative diseases with perturbation of these homeostatic systems. The process of SG/aggresome formation and clearance is important for human physiology such as antiviral responses and neurodegenerative diseases (Ramaswami et al., 2013). Mutations that increase SG formation or restrict SG clearance link to some neurodegenerative diseases, which is supported by genetic evidence. As mentioned above, neurodegenerative diseases (e.g. ALS/FTD)-related RNA binding proteins often have pathogenic mutations in their IDRs. For example, TIA1 harbors a LCS, where mutations are frequently observed in ALS/FTD, and these mutations alter the phase separation properties (Mackenzie et al., 2017). In cells, this perturbation of TIA1 results in more viscous SGs that fail to appropriately disassemble and persist long after the removal of stress. On the other hand, mutations of VCP/p97 chaperone protein, which is important for SG clearance, are observed in several neurodegenerative diseases including ALS and FTLD (Johnson et al., 2010). The disease-related mutants showed the constitutive accumulation in SGs, which could not be disassembled (Buchan et al., 2013). Thus, SGs are of interest for following two reasons. First, they are the representative RNP granules, suggesting that they reveal fundamental mechanism of RNA and protein compartmentalization through LLPS. Second, they would possibly reveal how to prevent pathogenic aggregation observed in neurodegenerative diseases.

1.4 Aim of this thesis

The laboratory of Prof. Dr. Patrick Matthias has focused on the physiological and pathological roles of HDACs. The aim of my Ph.D. thesis is to better understand the role of HDAC6. I aimed to specifically address and clarify the three questions below.

1. **How does HDAC6 deacetylate its substrate?**

   As I mentioned in **1.2.1**, the existence of tandem catalytic HDAC domains is unique to HDAC6 among HDAC family proteins. In parallel, HDAC6 has a specific preference for its substrates; it uniquely deacetylates α-tubulin. It thus has been assumed that this organization of two domains may confer the unique activity as α-tubulin deacetylase on HDAC6. Several studies tried to clarify how HDAC6 recognizes and deacetylates α-tubulin and other HDAC6-specific substrates, but there were not yet reasonable explanations mainly due to the lack of knowledge about structure of the catalytic domains. Although the crystal structure of the ZnF-UBP domain was originally reported six years ago (Ouyang et al., 2012), there was no available structure of the HDAC6 catalytic domains. Considering that HDAC6 plays important roles in various diseases, the structural information of its catalytic domains would be useful to accelerate drug development. Dr. Yasuyuki Miyake, a previous postdoc in the lab, tried to address this important question, and solved the structure of HDAC6 catalytic domains. I was also engaged in this project, and the outcomes are presented both in **2.2** (published work) and in **2.3** (appendix).

2. **How large population of proteins are targeted by HDAC6?**

   Besides lack of structural information and of the mechanism how HDAC6 deacetylate its substrates mentioned above, the repertoire of HDAC6 substrates itself was not enough to explain the versatile role of this enzyme in cellular processes. Even though there were several reported HDAC6 substrates, these were all identified individually in a specific setting (specific

cell types, experimental conditions etc.) and no studies provided us with a comprehensive view at proteome scale. In collaboration with the laboratory of Dr. Chunaram Choudhary (University of Copenhagen), we obtained HDAC6-dependent acetylome (presented in **2.1** and **2.3**). This analysis was effectively utilized to solve the next question 3, which was the central aim of my thesis.

3. How does HDAC6 control granule formation *in vivo*?

Previous work by Dr. SoHee Kwon, a former Ph.D. student in the lab revealed that HDAC6 is a critical component of stress granules (SGs). The authors demonstrated that the catalytic activity of HDAC6 is important for SG formation (Kwon et al., 2007), although there were no detailed molecular mechanisms reported. Considering that many proteins harboring IDR contributes to SG formation through LLPS, and also that PTMs can alter their LLPS propensities (discussed in **1.3.2**), we considered that HDAC6 may participate in this process by regulating cellular phase (liquid-liquid or liquid-solid) separation directly through deacetylation of IDR-containing substrates. The challenging point to tackle this question was the limited number of HDAC6 substrate; although there were some reports about HDAC6 substrates, none of the reported substrates had acetylation sites in their IDRs with LLPS propensity. Here, we utilized an HDAC6-dependent acetylome dataset mentioned above, identified novel substrates and discovered that HDAC6 preferentially targets acetylated lysines in IDRs. Among the newly identified HDAC6 substrates harboring acetylation sites in IDRs, we focused on the stress granule component "DEAD box RNA helicase 3, X-linked" (DDX3X), and demonstrated how HDAC6 contributes to SG formation with its catalytic activity. The outcomes related to this part are presented both in **2.1** (published work) and in **2.3** (appendix).

Results

2.1 Acetylation of intrinsically disordered regions regulates phase separation (Published, Nature Chemical Biology)

nature chemical biology

ARTICLES

https://doi.org/10.1038/s41589-018-0180-7

Acetylation of intrinsically disordered regions regulates phase separation

Makoto Saito[1,2], Daniel Hess[1], Jan Eglinger[1], Anatol W. Fritsch[3,4], Moritz Kreysing[3,4], Brian T. Weinert[5], Chunaram Choudhary[5] and Patrick Matthias[1,2*]

Liquid–liquid phase separation (LLPS) of proteins containing intrinsically disordered regions (IDRs) has been proposed as a mechanism underlying the formation of membrane-less organelles. Tight regulation of IDR behavior is essential to ensure that LLPS only takes place when necessary. Here, we report that IDR acetylation/deacetylation regulates LLPS and assembly of stress granules (SGs), membrane-less organelles forming in response to stress. Acetylome analysis revealed that the RNA helicase DDX3X, an important component of SGs, is a novel substrate of the deacetylase HDAC6. The N-terminal IDR of DDX3X (IDR1) can undergo LLPS in vitro, and its acetylation at multiple lysine residues impairs the formation of liquid droplets. We also demonstrated that enhanced LLPS propensity through deacetylation of DDX3X-IDR1 by HDAC6 is necessary for SG maturation, but not initiation. Our analysis provides a mechanistic framework to understand how acetylation and deacetylation of IDRs regulate LLPS spatiotemporally, and impact membrane-less organelle formation in vivo.

Cells contain a variety of membrane-less organelles that control diverse biological processes. It has recently been proposed that liquid–liquid phase separation (LLPS) driven by multivalent macromolecular interactions is a common mechanism contributing to the formation of membrane-less organelles[1,2]. Although the molecular determinants underlying LLPS in cells still remain unclear, several proteins harboring intrinsically disordered regions (IDRs) have been shown to phase separate under various solution conditions[3]. IDRs are determined by structural prediction analysis: they do not have defined 3D structure and often exhibit low complexity domains (LCDs) of repeated sequence elements that provide the basis for intermolecular interactions. A central question is that of mechanisms regulating the properties of IDRs in vivo to ensure that LLPS takes place when required.

Stress granules (SGs) and aggresomes are membrane-less organelles that arise when cells are under specific environmental stresses. SGs are composed of ribonucleoproteins (RNPs) and form when translation initiation is impaired; they represent a pool of mRNAs stalled in translation[4]. Aggresomes contain misfolded protein aggregates that are concentrated by microtubule-dependent motors to a perinuclear inclusion body and subsequently processed by autophagy[5]. Genetic evidence suggests that these granules are related to the inclusions seen in the pathogenesis of neurodegenerative diseases. Amyotrophic lateral sclerosis (ALS)-related proteins, such as FUS and hnRNPA1, undergo LLPS. Mutations within their IDRs are frequently associated with ALS, and these mutations alter their phase separation properties, including droplet and pathogenic fibrils formation in vitro and SG formation in vivo[6,7].

Post-translational modifications (PTMs) can alter the phase separation threshold of IDRs, and thus are recognized as important regulators for granule formation. They can change both the intrinsic solubility and valency[8]. For example, phase separation of the nuage protein DDX4 is hindered by arginine methylation in RGG motifs[9]. FUS harbors an N-terminal LCD and C-terminal domain (CTD) enriched in RGG/RG motifs, and both phosphorylation of the LCD and methylation in CTD inhibit its phase separation[10,11]. These findings confirm the effect of PTMs on phase separation, but their regulation and roles in a biological context are not clear.

The lysine deacetylase HDAC6 critically participates in the formation of SGs and aggresomes[12,13]. It localizes mainly to the cytoplasm, where α-tubulin is its most prominent substrate[14,15]. It has been implicated in various biological processes, including the stress response[16,17]. Under various stresses, HDAC6 co-localizes to SGs with the Ras-GTPase-activating protein SH3 domain-binding protein 1 (G3BP-1), an established component of SGs[18]. HDAC6 is important for the formation of SGs[13,19] as well as of the aggresome[12], although the contribution of HDAC6 catalytic activity to the process of granule formation is not understood. We hypothesized that HDAC6 may participate in this process by regulating cellular phase (liquid-liquid or liquid-solid) separation directly through substrate deacetylation.

Here, we used an HDAC6-dependent acetylome data set, and discovered that HDAC6 preferentially targets acetylated lysines in IDRs. Among these, we focused on the SG component "DEAD box RNA helicase 3, X-linked" (DDX3X). We found that oxidative and osmotic stress as well as translation inhibition elicit activation of the lysine acetyltransferase CBP, leading to acetylation of multiple proteins, including DDX3X, whereby several lysines in its N-terminal IDR (IDR1) are acetylated. Notably, purified DDX3X-IDR1 was sufficient to mediate droplet formation by LLPS in vitro, and this process was inhibited by acetylation. Impaired droplet formation in vitro was associated with the decreased total volume of SGs in vivo, indicating that acetylation of IDRs affects their LLPS and granule formation in a biological context. Deacetylation of DDX3X-IDR1 by HDAC6 is required for assembly of large, mature SGs. Mathematical modeling derived from live-cell imaging experiments demonstrates that acetylation does not impact the initial generation of SGs, but their maturation. These observations reveal that

[1]Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland. [2]Faculty of Sciences, University of Basel, Basel, Switzerland. [3]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany. [4]Center for Systems Biology Dresden, Dresden, Germany. [5]Department of Proteomics, The Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. *e-mail: patrick.matthias@fmi.ch

counteracting "writer" and "eraser" enzymes for PTMs modulate the LLPS propensity of IDR-containing proteins spatiotemporally in vivo, and that IDR lysine acetylation levels regulate membraneless organelle formation more broadly.

Results

HDAC6 preferentially targets acetylated lysines in IDRs. To determine whether acetylation and its regulation by HDAC6 may regulate LLPS, we re-analyzed acetylome data sets[20], comparing data obtained with two HDAC6-specific inhibitors (Tubacin or Bufexamac) to those obtained with HDAC6 knockout (KO) cells. About 4% of the quantified acetylated sites (87 out of 2029) displayed a more than two-fold increase in acetylation in at least one condition (Supplementary Fig. 1a) and are related to various cellular functions (Supplementary Fig. 1b). Moreover, about 1% of the sites (18 out of 2029) showed a more than two-fold increase in acetylation in all three conditions (Fig. 1a). These 18 high-confidence HDAC6 target sites are enriched with a GG motif immediately before the acetylated lysine GGK(Ac), which is targeted by CBP and p300 (ref. [21]; Supplementary Fig. 1c), and three sites are within the well-known HDAC6 substrate cortactin (Cttn), demonstrating the validity of the method[22] (Fig. 1a).

We next determined what fraction of the acetylated sites map to an IDR using two different structural prediction programs, VSL2 and VL3-BA; only 43–55% of the 2029 acetylated peptides fall within an IDR. In striking contrast, 83–94% of the high-confidence HDAC6 target sites map to an IDR (Fig. 1b and Supplementary Fig. 2), suggesting that HDAC6 may broadly regulate IDRs. Notably, previous work reported that acetylated lysines are significantly enriched in structured regions and depleted in unstructured regions[23]. Thus, the HDAC6 acetylome defines a skewed subset of the acetylated lysines, which may play a role in LLPS.

DDX3X is a substrate of both HDAC6 catalytic domains. Among the high-confidence HDAC6 target sites, DDX3X-K118 acetylation (SGFGK(Ac)FER) caught our interest because it showed one of the highest SILAC ratios in our data sets. DDX3X harbors two IDRs, one at the N terminus, where K118 is localized, and one at the C terminus (Fig. 1c). DDX3X was also independently reported to be acetylated in mouse liver[24] (Supplementary Fig. 3), indicating that the regulation of DDX3X by acetylation is also observed in the whole organism. Furthermore, HDAC6 and DDX3X were both independently reported to be localized to SGs under oxidative stress[13,25], offering a possible link between HDAC6 and DDX3X in the process of granule formation (Supplementary Fig. 4). We first evaluated the acetylation status of DDX3X in cells to confirm the data from the acetylome analysis.

To test which histone acetyltransferases (HATs) can act on DDX3X, we expressed four different HATs in HEK293T (hereafter 293T) cells and detected acetylation of DDX3X. Both CBP and p300, but not PCAF or Tip60, led to acetylation of endogenous or overexpressed DDX3X (CBP: 3.8 ± 0.5, p300: 2.0 ± 0.4, PCAF: 1.1 ± 0.1, and Tip60: 1.2 ± 0.2 fold; Fig. 1d; Supplementary Figs. 5a and 6a). HDAC6 inhibition and acetylation by CBP showed a cooperative effect on the acetylation level of overexpressed DDX3X (Supplementary Fig. 5b) as well as of endogenous DDX3X (Tubacin: 1.2 ± 0.03 fold, CBP: 2.0 ± 0.2 fold, and CBP + Tubacin: 4.0 ± 0.9 fold; Fig. 1e and Supplementary Fig. 6b). To confirm that DDX3X-K118 is acetylated by CBP, we co-expressed CBP together with a DDX3X-K118R mutant, which cannot be acetylated, in 293T cells. DDX3X-K118R displayed a weaker signal compared to DDX3X-WT, indicating that K118 is a major acetylation site but that there are also other acetylation sites (Supplementary Fig. 5c); indeed, four additional acetylation sites were detected in the IDR1 (K50, K64, K66, and K81; Supplementary Fig. 5d,e). This result is consistent with another acetylome data set, which revealed that several lysines in IDR1 (K50, K55, K64, K66, K118 and K130) are acetylated[20]. As K118 exhibited the highest SILAC ratio among identified sites, we generated a DDX3X-K118Ac-specific antibody (Supplementary Fig. 7); in the following experiments this antibody was used to monitor K118Ac as an indicator of IDR1 acetylation.

Different deacetylases often redundantly target the same lysine acetylation site on a target protein[14,22,26]. To test whether deacetylation of DDX3X-K118Ac is specific for HDAC6, we transiently expressed CBP in 293T cells treated with several HDAC inhibitors and assessed DDX3X-K118Ac. DDX3X-K118Ac increased upon treatment with the HDAC6-specific inhibitors, Bufexamac (3.5 ± 0.7 fold) and Tubacin (4.4 ± 0.9 fold) (Fig. 1f and Supplementary Fig. 6c). Moreover, Tubastatin A, another HDAC6 specific inhibitor, also increased DDX3X-K118Ac (2.6 ± 0.5 fold), further supporting the notion that this DDX3X residue is an HDAC6 target site. In contrast, neither nicotinamide, a sirtuin deacetylase inhibitor, nor sodium butyrate, a class I HDAC inhibitor, increased the acetylation status of DDX3X-K118 (Nicotinamide: 0.7 ± 0.05, Sodium butyrate: 0.3 ± 0.1 fold). To define the domains of HDAC6 required for DDX3X deacetylation, we established HDAC6 KO 293T cells (Supplementary Fig. 8) and used them to express various HDAC6 functional mutants (H216A, CD1 catalytic dead; H611A, CD2 catalytic dead; H216A/H611A double mutant, and W1182A, ubiquitin-binding-deficient)[27,28]. We monitored the acetylation level of DDX3X-K118, α-tubulin-K40, and cortactin. Each single catalytic domain HDAC6 mutant could robustly deacetylate DDX3X (Ac-DDX3X signal; WT: 0.1 ± 0.02, H216A: 0.1 ± 0.04, H611A: 0.3 ± 0.05, H216A/H611A: 0.9 ± 0.1, and W1182A: 0.3 ± 0.03 fold), and also to a lesser extent cortactin, but not α-tubulin (Ac-α-tubulin signal; WT: 0.5 ± 0.1, H216A: 0.5 ± 0.04, H611A: 1.0 ± 0.1, H216A/H611A: 1.1 ± 0.1, and W1182A: 0.7 ± 0.1 fold; Supplementary Fig. 9). This indicates that either HDAC6 domain, CD1 or CD2, can deacetylate DDX3X or cortactin, two proteins containing an acetylated IDR. This reactivity is different from the reactivity for α-tubulin, where the acetylated residue is in a structured region: deacetylation of α-tubulin requires the catalytic activity from CD2 (refs [29,30]; Supplementary Fig. 9). Together, these results indicate that DDX3X-K118Ac within IDR1 is a specific target of HDAC6 and also suggest that CD1 may be generally active on acetylated lysines in IDRs, which has implications for the generation of novel inhibitors.

DDX3X is acetylated under specific stress conditions. We hypothesized that deacetylation of DDX3X-IDR1 may be regulated by HDAC6 during the stress response process; we performed a time-course of arsenite treatment and compared 293T and HDAC6 KO 293T cells lysates by immunoblotting (Fig. 2a). Phosphorylation of eIF2α, which identifies stress-induced translational arrest, was observed from 10 min onward after arsenite treatment. We found that in 293T cells acetylation of DDX3X-IDR1 is biphasic: it gradually increased in the early phase (0 to 45 min), and decreased later (from 45 to 60 min; Fig. 2a, lanes –5). In contrast, in HDAC6 KO 293T cells acetylation of DDX3X continued to increase between 45 min and 60 min (Fig. 2a, lanes 6–10; Fig. 2b). This result is consistent with the fact that DDX3X and HDAC6 localize to SGs, whereas CBP, which is mostly nuclear, shows a weak cytoplasmic staining outside of SGs during the whole process (Supplementary Fig. 10a), and suggests that HDAC6 effectively deacetylates DDX3X in SGs.

To define the physiological stresses where acetylation of DDX3X is enhanced, we systematically treated 293T cells with inducers for various types of stress: oxidative stress (sodium arsenite, H_2O_2 and diethyl maleate), energy depletion (CCCP and clotrimazole), translation inhibition (puromycin), ER stress (thapsigargin), proteasome inhibition (MG132), heat shock and osmotic stress (sorbitol), and confirmed SG formation under each condition (Supplementary Fig. 10b); phosphorylation of eIF2α was used to determine the best time points for each inducer. The same experiment was then

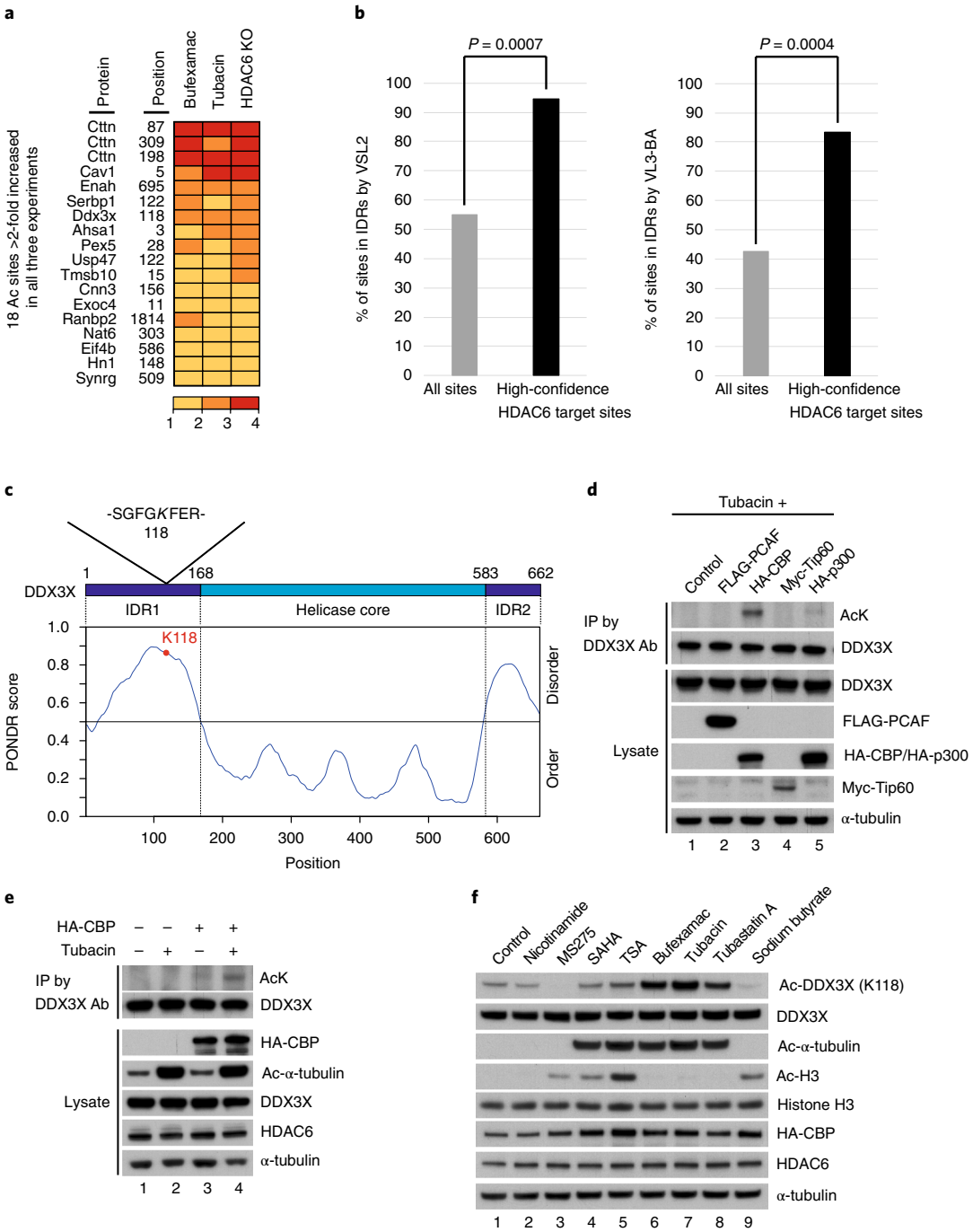


Fig. 1 | DDX3X-IDR1 is specifically deacetylated by HDAC6. a, The heatmap shows the quantified 18 high-confidence peptides upregulated two-fold or more in all three conditions (based on the data set from ref. [20]). **b**, Almost all of the 18 high-confidence HDAC6 target sites map to IDRs. Structural prediction analysis was performed with the 18 high-confidence HDAC6 target sites as well as with all (2,029) acetylated sites examined here, as a control. The VSL2 and VL3-BA programs of the PONDR suite (Prediction Of Naturally Disordered Regions) were used, as indicated. *P* values were determined by chi-squared test; $***P < 0.001$. **c**, Analysis of the DDX3X protein sequence for intrinsically disordered regions, using the program VL3-BA. The amino acid position in the linear sequence is plotted on the *x* axis, and the disorder tendency is displayed on the *y* axis. Both the N-terminal and the C-terminal regions of DDX3X show a strong disorder potential. Notably, residue K118 is found in the region with the maximum disorder potential. **d**, Acetylation of endogenous DDX3X by CBP and p300. Tagged versions of PCAF, CBP, Tip60 and p300 were transiently transfected into 293T cells. The acetylation status of endogenous DDX3X under Tubacin was analyzed by immunoblotting with a pan-acetyllysine antibody (AcK) following DDX3X immunoprecipitation (IP). Representative results are shown ($n = 3$ biologically independent samples). Uncropped gel images can be found in Supplementary Fig. 27a. **e**, Enhanced acetylation of endogenous DDX3X following Tubacin treatment. 293T cells expressing CBP were treated with 10 μM Tubacin for 12 h. The acetylation status of DDX3X was analyzed with a pan-acetyllysine antibody following immunoprecipitation. Representative results are shown ($n = 3$ biologically independent samples). Uncropped gel images can be found in Supplementary Fig. 27b. **f**, Effects of HDAC inhibitors on DDX3X-K118Ac in 293T cells. CBP was transiently expressed in 293T cells treated overnight with different HDAC inhibitors, as indicated. The acetylation status of endogenous DDX3X was tested with the DDX3X-K118Ac-specific antibody. Representative results are shown ($n = 3$ biologically independent samples). Uncropped gel images can be found in Supplementary Fig. 27c.

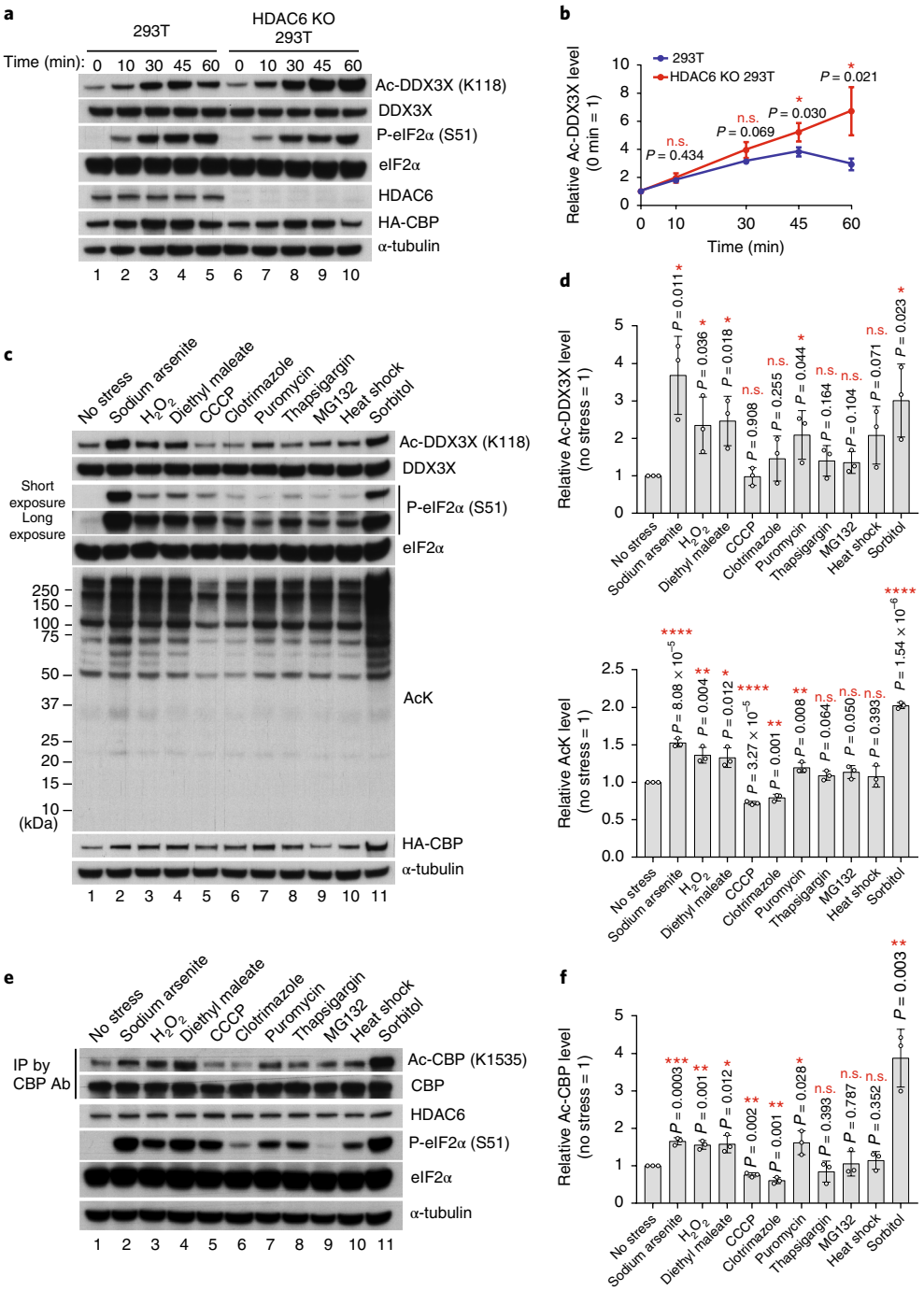


Fig. 2 | Stress induces acetylation of DDX3X and other proteins. a, DDX3X-K118 acetylation in 293T and HDAC6 KO 293T cells under arsenite treatment. The acetylation status of DDX3X was analyzed in both 293T and HDAC6 KO 293T cells after CBP expression and 1 mM arsenite treatment for the indicated times. Representative results are shown ($n = 3$ biologically independent samples). Uncropped gel images can be found in Supplementary Fig. 27d. **b**, Quantification of Ac-DDX3X signals as in **a**. Intensity of the band at each time point was quantified using ImageJ, and normalized by the signal at 0 min. (mean ± s.d., $n = 3$ biologically independent samples). **c**, DDX3X-IDR1 acetylation increases under various types of stress. HDAC6 KO 293T cells expressing CBP were treated with different stress inducers (1 mM sodium arsenite for 1 h, 2 mM H_2O_2 for 1 h, 3 mM diethyl maleate for 1 h, 20 μM CCCP in glucose-free culture medium for 1 h, 40 μM clotrimazole in glucose-free culture medium for 3 h, 20 μg/ml puromycin for 3 h, 10 μM thapsigargin for 1 h, 10 μM MG132 for 3 h, heat shock at 43 °C for 1 h and 0.5 M sorbitol for 1 h). Cell cytosolic extracts were analyzed by immunoblotting with the indicated antibodies. eIF2α phosphorylation was used as a marker of cellular stress. Representative results are shown ($n = 3$ biologically independent samples). Uncropped gel images can be found in Supplementary Fig. 27e. **d**, Quantification of Ac-DDX3X signals and AcK signals in **c** (mean ± s.d.; $n = 3$ biologically independent samples). **e**, CBP acetylation is induced by various types of stress. WT MEF cells were treated with different stress inducers and analyzed as in **d** above. Representative results are shown ($n = 3$ biologically independent samples). Uncropped gel images can be found in Supplementary Fig. 27f. **f**, Quantification of Ac-CBP signals in **e** (mean ± s.d.; $n = 3$ biologically independent samples). *P* values were determined by Student's two-tailed *t*-test; $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ and $^{****}P < 0.0001$.

repeated with HDAC6 KO 293T cells, and acetylation of DDX3X-IDR1 was monitored (Fig. 2c). As DDX3X is an HDAC6 specific substrate, its acetylation is expected to be stable in cells lacking HDAC6. Oxidative stress inducers including arsenite, puromycin and sorbitol all significantly increased acetylation of DDX3X-K118 (arsenite: 3.7 ± 0.9, H_2O_2: 2.4 ± 0.6, diethyl maleate: 2.5 ± 0.5, puromycin: 2.1 ± 0.5, sorbitol: 3.0 ± 0.8 fold; Fig. 2d). In contrast, acetylation of DDX3X was not significantly altered under any other condition, although phosphorylation of eIF2α was observed (Fig. 2c). Global protein acetylation was observed with the same inducers (arsenite: 1.5 ± 0.05, H_2O_2: 1.4 ± 0.1, diethyl maleate: 1.3 ± 0.1, puromycin: 1.2 ± 0.1, sorbitol: 2.0 ± 0.03 fold; Fig. 2d) as for DDX3X acetylation (Fig. 2c,d). We next tested whether stress induction might lead to elevated CBP activity. CBP autoacetylates and its level of acetylation can be taken as a measure of its activity[31]. We treated mouse embryonic fibroblasts (MEFs) with different stress inducers and examined the level of endogenous CBP acetylation (Fig. 2e). As seen in Fig. 2f, oxidative stress, puromycin and sorbitol lead to a robust increase in CBP acetylation (arsenite: 1.7 ± 0.1, H_2O_2: 1.6 ± 0.1, diethyl maleate: 1.6 ± 0.2, puromycin: 1.6 ± 0.3, sorbitol: 3.9 ± 0.6 fold), thus establishing a direct link between the stress response and protein acetylation.

Acetylation of DDX3X-IDR1 impairs its LLPS. We next investigated the role of DDX3X-IDR1 acetylation in LLPS; for this, we expressed and purified the IDR1 (Supplementary Fig. 11a). When polyethylene glycol (PEG) was added to the sample, it became turbid, and micron-sized droplets were observed by differential interference contrast (DIC) microscopy (Supplementary Fig. 11b). Droplets of IDR1 fused to each other (Supplementary Video 1) and were also sensitive to salt concentration (Supplementary Fig. 11c) and temperature (Supplementary Video 2 and Supplementary Fig. 11d), indicating that IDR1 undergoes LLPS.

To monitor the effect of DDX3X-IDR1 acetylation on LLPS, the turbidity (OD_{600}) of unacetylated and acetylated IDR1 during LLPS was quantified. Incubation of purified IDR1 with CBP and acetyl-CoA led to acetylation of all of its ten lysine residues, including K118 (Supplementary Fig. 11e). Notably, the turbidity of this acetylated IDR1 solution was much lower compared to the solution of unacetylated IDR1 (Fig. 3a). Moreover, the turbidity was restored upon co-incubation of acetylated IDR1 with the purified catalytic domains of HDAC6 (ref. [30]; Fig. 3b), indicating that lysine acetylation affects LLPS. We hypothesized that this effect is due to the loss of intermolecular interactions between cationic lysine and anion or aromatic π, caused by the neutralization of positively charged lysine by acetylation. Therefore, we further confirmed this effect using acetyllysine mimic (K to Q) mutants. We expressed and purified K118Q, K118R and allQ (all IDR1 lysines mutated to glutamine) mutants (Supplementary Fig. 11f) and measured their turbidity during LLPS. The allQ mutant showed the lowest turbidity among the samples, and this effect was clearly visible with its transparent solution under conditions wherein the WT solution was opalescent (Fig. 3c). The K118Q mutant also showed lower turbidity compared to WT; in contrast, this effect was not observed with the K118R mutant, which maintains the positive charge (Fig. 3c). These results demonstrate that DDX3X-IDR1 plays a role in LLPS, and reveal that a single acetylation on K118 is already sufficient to robustly alter the dynamics of droplet formation in vitro.

Although our data about the acetyl-mimic K118Q mutant indicate that acetylation on a single lysine residue is sufficient to affect LLPS, acetylation on multiple lysine residues (mimicked by allQ mutant) shows a stronger effect. This indicates the possibility that each lysine within the IDR1 may contribute to efficient LLPS. To test this, we compared the LLPS efficiency of individual K to Q mutants (Supplementary Fig. 12a) to that of WT and allQ mutant. Seven mutants (K35, 50, 55, 64, 66, 81 and 118) showed significantly lower turbidity compared to WT, but none of them showed the severe impairment as the allQ mutant did (Fig. 3d and Supplementary Fig. 12b). In contrast, three mutants (K130, 138 and 162) formed droplets to the same extent as the WT. One study may explain these positional effects on LLPS: the DDX3X region (a.a. 1-167) that is predicted to be an IDR by several established algorithms has a conserved region (a.a. 132-167 in human and mouse) which harbors a predicted short α-helix[32]. Although secondary structures such as α -helix were also reported to trigger LLPS[33], our observations indicate that in the case of DDX3X, lysine acetylation in a 'genuine' IDR impairs LLPS more effectively than acetylation in an α-helix.

Next, we investigated LLPS of purified full-length DDX3X (Supplementary Fig. 13a). The concentration of cytosolic DDX3X was estimated to be around 0.8 μM (Supplementary Fig. 13b), and we confirmed that acetyllysine mimic mutants (K118Q and allQ) at this physiological concentration still exhibit less turbidity compared to WT in droplet-formation assays (Supplementary Fig. 13c). This effect was also observed with WT DDX3X acetylated by CBP, and counteracted by the catalytic domains of HDAC6, as initially shown with the IDR1 (Supplementary Fig. 13d). Similarly, we also observed that the IDR of another HDAC6 substrate, cortactin[34], undergoes LLPS. In this case as well, acetyl-mimic mutation impairs LLPS (Supplementary Fig. 14). Thus, our observations with DDX3X hint to a general mechanism for LLPS regulation by acetylation.

To further examine the role of acetylation on the capacity of DDX3X-IDR1 to assemble in granules, we performed *in vitro* experiments with the chemical biotinylated isoxazole (b-isox), that can selectively precipitate proteins with LCDs from cellular lysates[35]. We found that b-isox precipitated DDX3X and another LCD-containing protein, TIAR (Fig. 3e, lanes 3 and 4), but not acetylated DDX3X (lane 4, compare with lane 2), demonstrating that the IDR1 acetylation status is critical for DDX3X to form crystals with b-isox. Taken together, these experiments revealed that DDX3X undergoes LLPS, which is impaired by acetylation-mediated neutralization of cationic lysines in the IDR1.

DDX3X-IDR1 is important for SG formation. The above findings about LLPS of DDX3X regulated by acetylation support our hypothesis that DDX3X is one of the effector substrates regulated by HDAC6 during SG formation. To confirm liquid-like properties of DDX3X-positive SGs, an mCherry-DDX3X fusion protein was expressed in MEFs, and the dynamics of SG formation were monitored over time. We observed that granules fusion is accompanied by first a transient change of overall shape, which is followed by gradual recovery of the initial round shape; this behavior has the properties expected from LLPS (see examples in Supplementary Video 3 and Supplementary Fig. 15). Although the role of DDX3X in SGs had already been examined[25], its stress dependence was unclear. To address this question, we generated DDX3X KO MEFs (Supplementary Fig. 16), and SG formation was monitored in DDX3X KO and parental MEFs under various types of stresses (Supplementary Fig. 17a). We quantified the total area of G3BP foci, as done by ref. [36], and observed clear defects in G3BP granule formation under oxidative stress, energy depletion and translation inhibition (Supplementary Fig. 17b), indicating that DDX3X is important for SG assembly under these specific stress conditions. The observed impairment was not due to changes in expression level of G3BP or phosphorylation level of eIF2α (Supplementary Fig. 17c,d).

We anticipated that the DDX3X-IDR1 would contribute to LLPS in cells and thus might impact SG formation; to test this directly, we expressed DDX3X mutants lacking different domains in HeLa cells and monitored their behavior in SG localization (Supplementary Fig. 18). Although the DDX3X helicase core itself does not localize to SGs under arsenite treatment, the construct including the IDR1 (del-IDR2) does (Supplementary Fig. 18c). It is important to note that the IDR1 is necessary but not sufficient for SG localization;

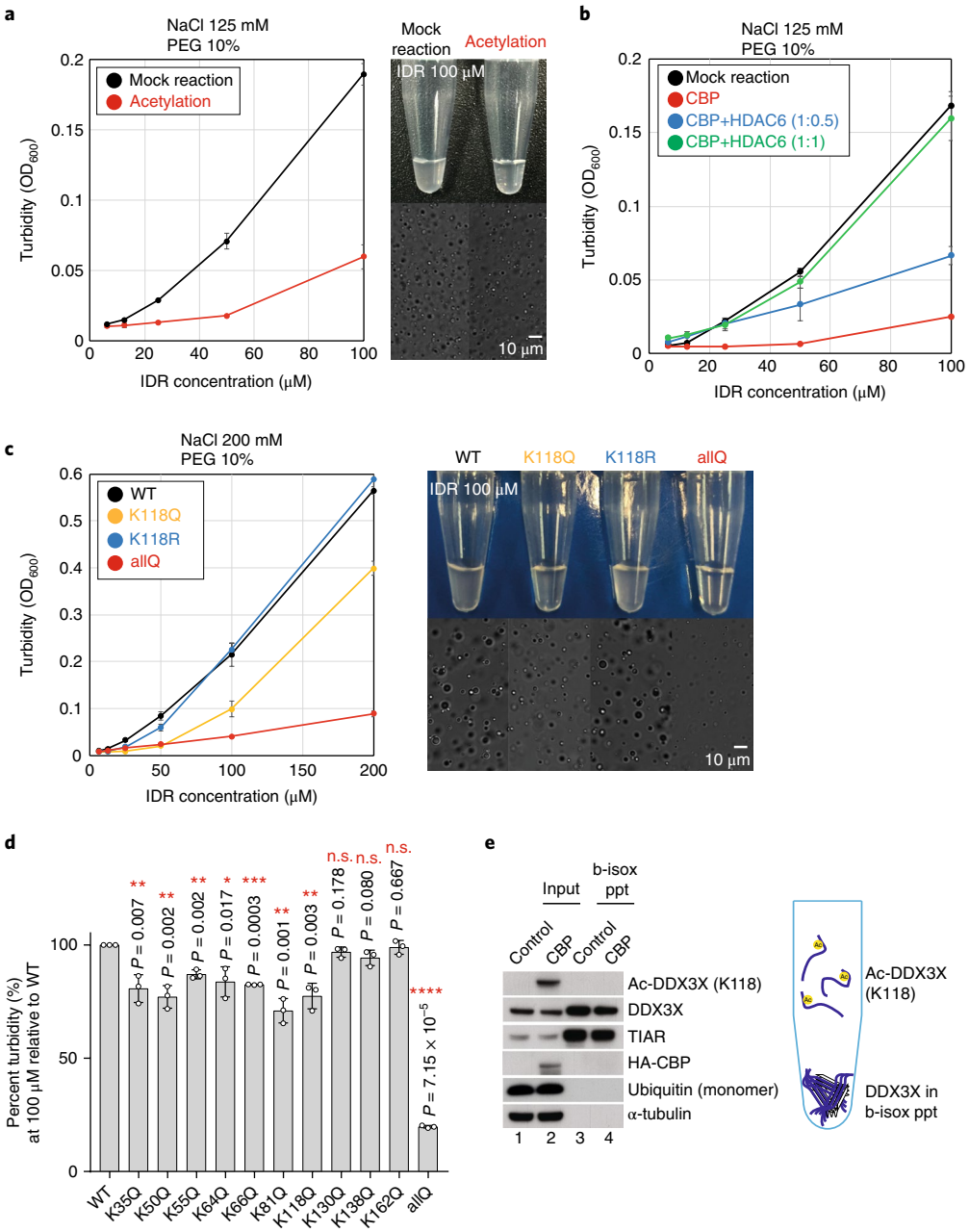


Fig. 3 | Acetylation of DDX3X-IDR1 impairs its droplet formation by LLPS in vitro. a, Effect of acetylation on LLPS of DDX3X-IDR1. Left, turbidity (OD_{600}) of DDX3X-IDR1 acetylated by CBP (acetylation) and the mock reaction control (mock reaction) in 20 mM Tris, pH 7.5, 125 mM NaCl, 2 mM DTT and 10% PEG was measured at room temperature (mean ± s.d.; $n = 3$ independent experiments). For the mock reaction control, the IDR1, CBP and acetyl-CoA were mixed directly before the turbidity measurement. Right, phase separation of 100 μM IDR1 solutions (top). Liquid droplets visualized by DIC microscopy (bottom; scale bar, 10 μm). **b**, Recovery of LLPS by HDAC6-mediated deacetylation. Turbidity (OD_{600}) of DDX3X-IDR1 was measured after the reaction with CBP or CBP + HDAC6 (mean ± s.d.; $n = 3$ independent experiments). The HDAC6 reaction was carried out at two concentrations (molar ratio between CBP and HDAC6 = 1:0.5 (blue), 1:1 (green)). **c**, Left, lysine to glutamine (acetyllysine mimic) substitution impairs LLPS of DDX3X-IDR1. Turbidity (OD_{600}) of WT, K118Q, K118R and allQ mutants of DDX3X-IDR1 in 20 mM Tris, pH 7.5, 200 mM NaCl, 2 mM DTT and 10% PEG was measured (mean ± s.d.; $n = 3$ independent experiments). Right, phase separation of 100 μM IDR1 solutions (top) and liquid droplets visualized by DIC microscopy (bottom; scale bar, 10 μm). **d**, Percent of turbidity (OD_{600}) of DDX3X-IDR1 mutants from Supplementary Fig. 12b, normalized to WT (mean ± s.d.; $n = 3$ independent experiments). *P* values were determined by Student's two-tailed *t*-test; $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ and $^{****}P < 0.0001$. **e**, Acetylation of the DDX3X IDR regulates its ability to precipitate with biotinylated isoxazole (b-isox) in vitro. Left, lysates from 293T cells expressing GFP (control) or CBP were subjected to b-isox-mediated precipitation and analyzed by immunoblotting. The presence of endogenous DDX3X-K118Ac was assessed in the precipitate fraction (b-isox ppt) and in the lysate fraction (input) and compared to the level of total precipitated DDX3X. TIAR was used as a control protein harboring an IDR, whereas ubiquitin and α-tubulin were used as control proteins lacking an IDR. Proteins were detected with specific antibodies, as indicated. Representative results are shown ($n = 3$ biologically independent samples). Uncropped gel images can be found in Supplementary Fig. 27g. Right, schematic illustrating the b-isox-mediated precipitation of unacetylated DDX3X IDRs with acetylated IDRs remaining in suspension.

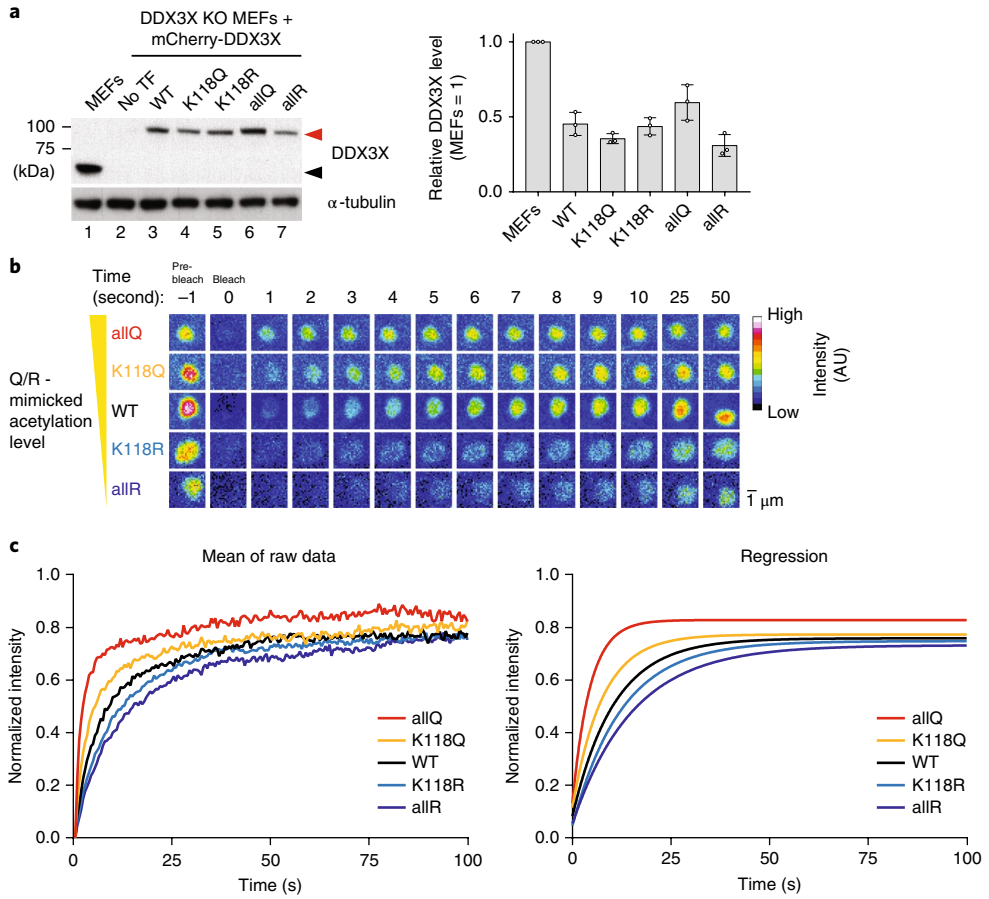


Fig. 4 | Acetyl-mimic/dead mutations alter DDX3X SG dynamics. a, Transient expression of mCherry-DDX3X WT and acetyl-mimic/dead mutants in DDX3X KO MEFs. Left, the expression level of DDX3X 24 h after transfection was analyzed by immunoblotting; the black arrowhead indicates endogenous DDX3X, and the red one indicates mCherry-DDX3X. Right, quantification of DDX3X signal (mean ± s.d.; $n = 3$ biologically independent samples). Uncropped gel images can be found in Supplementary Fig. 27h. **b**, Fluorescence recovery after photobleaching (FRAP) of mCherry-DDX3X SGs. Representative images of time-dependent fluorescence recovery are shown. Cells at 24 h after transfection were treated with 1 mM arsenite for 30 min. SGs with each mCherry-DDX3X construct were photobleached (at 0 s), and their recovery was monitored in 1 frame/second. Scale bar, 1 μm. **c**, Quantification of results in **b** ($n = 5$ biologically independent experiments). FRAP curves after photobleaching of each mCherry-DDX3X SG are shown (left: raw data; right: regression curves).

thus, both the IDR1 and the helicase core interacting with RNAs contribute to SG localization, which is consistent with a molecular feature increasingly observed in RNPs. Furthermore, the reduced total area of G3BP foci in cells lacking DDX3X is restored to WT levels by re-expression of WT or del-IDR2 DDX3X (Supplementary Fig. 18d). This observation identifies the critical role of the IDR1 for SG formation and encouraged us to investigate the role of its acetylation in this process.

Deacetylation of DDX3X-IDR1 is required for SG formation. To investigate acetylation of DDX3X-IDR1 in SG formation, we first used fluorescence recovery after photobleaching (FRAP) to measure the mobility of mCherry-DDX3X. For this we used DDX3X KO MEFs and transiently re-expressed mCherry-DDX3X WT or mutants thereof: two acetyl-mimic mutants (K118Q, allQ) and two acetyl-dead mutants (K118R, allR). The expression level of each protein was comparable to that of endogenous DDX3X (Fig. 4a). Furthermore, the allR mutant (which is fully unacetylated, but preserves positive charges) served as a control to monitor the acetylation status of other constructs under physiological conditions. All DDX3X constructs formed SGs under arsenite, and we found that acetyl-mimic or dead mutations significantly altered the dynamic exchange of DDX3X inside and outside SGs (Fig. 4b). Compared to WT, acetyl-mimic mutants exhibited decreased half-recovery times and a larger mobile fraction (Fig. 4c and Supplementary Fig. 19), suggesting that their granules are less viscous than those made by WT. In contrast, granules of acetyl-dead mutants were more viscous (Fig. 4c and Supplementary Fig. 19). Importantly, these effects reflect the degree of K to Q or K to R mutations; allQ and allR mutants have the most significant effects, whereas K118Q and K118R mutants are mildly affected (Fig. 4c and Supplementary Fig. 19). The clear phenotypic difference between WT and allR DDX3X provides evidence for DDX3X acetylation under physiological conditions. To confirm this, we established rescue lines stably expressing either DDX3X WT or each mutant and visualized Ac-DDX3X by immunofluorescence microscopy. The expression level of each construct was comparable with that of endogenous DDX3X (Supplementary Fig. 20a), and broadly distributed signals of Ac-DDX3X in cytosol including SGs were observed with WT, but not with allR rescuant (Supplementary Fig. 20b).

To test whether the impaired LLPS propensity of acetyl-mimic mutants has an impact on SG formation, we quantified the volume of SGs under arsenite: we found that the total volume of SGs per cell decreased in both acetyl-mimic mutants, but increased in the acetyl-dead allR mutant (Fig. 5a,b and Supplementary Dataset 1). In agreement with this, acetylation of endogenous DDX3X induced

by ectopic expression of CBP also decreased the total volume of SGs in WT MEF cells (Supplementary Fig. 21 and Supplementary Dataset 1). Given these results, we next investigated whether endogenous CBP and HDAC6 counteract each other to regulate the acetylation status of DDX3X-IDR1 and SG formation. After pretreatment with an inhibitor of p300/CBP (A-485) or HDAC6 (Tubacin), we treated DDX3X WT rescuant cells with five stress inducers, which lead to increased CBP autoacetylation and DDX3X acetylation (Supplementary Fig. 22). As expected, inhibition of CBP increased the total volume of SGs, and inhibition of HDAC6 decreased it under four stress inducers (arsenite, H_2O_2, diethyl maleate, puromycin) in which DDX3X is important for SG formation (Fig. 5c and Supplementary Dataset 1). This effect was not observed under sorbitol, as DDX3X does not contribute to SG formation there (Fig. 5c and Supplementary Dataset 1). We also tested the SG volume of each mutant under arsenite in combination with A-485 or Tubacin and noticed that the effects of these inhibitors are mitigated (Fig. 5d and Supplementary Dataset 1), suggesting that CBP and HDAC6 regulate SG formation by modulating IDR1 acetylation. Consistent with this pharmacological inhibition, knockdown of CBP also led to a similar result (Supplementary Fig. 23 and Supplementary Dataset 1). We also monitored the disappearance of SGs over time. After 1 h recovery, about 77% of the cells expressing WT still had detectable SGs, but only 26% of the cells expressing allQ mutant or 54% of the cells expressing K118Q mutant did (Fig. 5e). Together, these experiments indicate that acetylation of IDR1 impairs the formation and stability of DDX3X SGs and that this is regulated by CBP and HDAC6 under physiological conditions. Our earlier work had demonstrated that loss of HDAC6 reduces the number of large SGs under arsenite treatment[13], which is likely related to the acetylation of DDX3X shown here (Supplementary Fig. 24). Consistent with this hypothesis, inhibition of CBP or HDAC6 does not affect SG volume when the DDX3X-IDR1 is mutated.

SG maturation is promoted by deacetylation of DDX3X. SG formation has recently been proposed as a two-step process, with initial nucleation of a stable'core' followed by maturation with accumulation of proteins harboring IDRs into a peripheral 'shell' through LLPS[6,37–39]. Many components of SGs can undergo LLPS, and heterotypic interactions among these proteins contribute to SG maturation. Given this model and our observation that HDAC6 deacetylates DDX3X to recover its LLPS propensity in later stage of SG formation (Fig. 2a,b), we hypothesized that LLPS of DDX3X is important for inducing LLPS of other SG components as well as resulting SG maturation. To examine this hypothesis, we immunoprecipitated DDX3X WT and allQ/R mutants from arsenite-treated cells, and their interactomes were defined (Supplementary Fig. 25a and Supplementary Dataset 2). The DDX3X interactomes are consistent with SG components reported in recent studies[37,40]. Strikingly, the comparison between WT and allQ/R mutants revealed that several SG proteins are not enriched with the allQ mutant, while they interact with the WT and allR mutant: HNRNPA2/B1 and DHX9, which are harboring GY-rich IDRs, and PABPC1, whose homolog in *S. cerevisiae* (Pab1) undergoes LLPS[41] (Supplementary Fig. 25b). This result indicates that interactions between cationic lysines in the DDX3X IDR and aromatic π in IDRs are impaired in the allQ mutant. Moreover, this suggests that the LLPS propensity of DDX3X is important for recruitment of additional proteins which themselves can undergo LLPS and contribute to the formation of the SG shell.

In addition, we also analyzed the growth of SGs formed by different DDX3X mutants (Supplementary Video 4), followed by mathematical modeling. Remarkably, the data (Supplementary Dataset 3) is well described by the Avrami equation for nucleated particle growth or film growth[42].

$$A(t) = A_F(1-e^{-k(t-T)})$$

This simple model depends only on three parameters: the total area of the SGs, A_F, at steady state, their growth rate, k, and the initiation time of SGs T (measured with respect to the start time of arsenite treatment, set to $t=0$). Because T can be determined directly from the time-course experiment, only A_F and k need to be fitted. This model has recently been used to model the growth of nuclear DDX4 organelles in cells[9]. For each DDX3X protein SG growth was observed at roughly the same time after arsenite addition (WT: 13.2 ± 2.3 min, K118Q: 13.5 ± 3.9 min, and allQ: 16.4 ± 3.1 min; Supplementary Fig. 26a). The result from five cells expressing each mutant is plotted (Supplementary Fig. 26b). After normalization of A_F by A_0 (cell area at $t=0$, see Supplementary Fig. 26c), the parameters for fitting of WT and mutants (each $n=5$) were compared (Supplementary Fig. 26a). Fitting the growth of WT DDX3X-positive SGs over time to this model resulted in a characteristic growth rate of $k=0.11\pm0.014\,\mathrm{min}^{-1}$ with a scaling ratio r ($=A_F/A_0$) of 10.1 ± 3.3 ($\times10^{-3}$). The kinetics of the acetyl-mimic K118Q DDX3X-positive SGs also fit with the equation above, and k ($0.12\pm0.061\,\mathrm{min}^{-1}$) is similar to that of WT. However, the ratio r is significantly different (K118Q: 4.0 ± 1.6 ($\times10^{-3}$)), indicating a reduced steady-state level. The allQ mutant also shows significant difference in the ratio r with an even much lower steady-state level (1.4 ± 0.6 ($\times10^{-3}$); Fig. 6a and Supplementary Fig. 26a). Together, these results suggest that granule growth of the acetyl-mimic mutants starts as for the WT protein, but becomes impaired during the process due to low LLPS efficiency; deacetylation by HDAC6 is then required for SG maturation (Fig. 6b).

Discussion

The central concept revealed by our study is that lysine acetylation in IDRs regulates LLPS, and impacts SG maturation via DDX3X, a novel HDAC6 substrate. Cancer-associated mutations in DDX3X map to its helicase domain and correlate with an altered tendency to form SGs[43]. LAF-1, a related *C. elegans* protein, had been shown to undergo LLPS[44], but the contribution of the DDX3X-IDR1 to LLPS had not been examined.

A few studies have suggested a role for lysine acetylation on granule formation under stress, although without mechanistic insight. An RNAi screen identified the HAT p300 as an important factor for SG formation[45], while another study described the deacetylase SIRT6 as a critical factor[46]; these two findings, indicating that both HATs and HDACs are implicated, underscore the role of lysine acetylation for SG formation. We now demonstrate that spatiotemporal acetylation of DDX3X-IDR1, regulated by CBP and HDAC6, is important for controlling SG formation. Specific stresses induce autoacetylation and activation of CBP, and lead to global protein acetylation, including DDX3X. Deacetylation of DDX3X by HDAC6 is then necessary for robust LLPS and SG formation. Our results show that this process is critical not for the initiation of the SGs, but for their maturation, which is consistent with the recent "Cores First" model[4]. Another RNP component, TDP-43, also becomes acetylated in its RNA-binding domain by CBP under oxidative stress. In this case, acetylation impairs its RNA binding ability and aberrant aggregate formation is observed, which is counteracted by HDAC6[47]. Thus, acetylation impacts granules formation in different ways, by modulating the LLPS propensity (DDX3X) or by affecting the ability to bind RNA (TDP-43); in each case, the action of HDAC6 is required for normal resolution of the stress response. It is currently unknown how specific stresses lead to activation of CBP with concomitant acetylation of multiple cellular targets, how the LLPS propensity of different acetylated proteins is altered, and whether or not this global acetylation provides any benefits to cellular fitness. However, active deacetylation by HDAC6 of some RNP components (DDX3X, TDP-43, others?) is essential for their normal dynamics and function in the stress response. Thus, our results demonstrate a physiological role for acetylation/deacetylation in

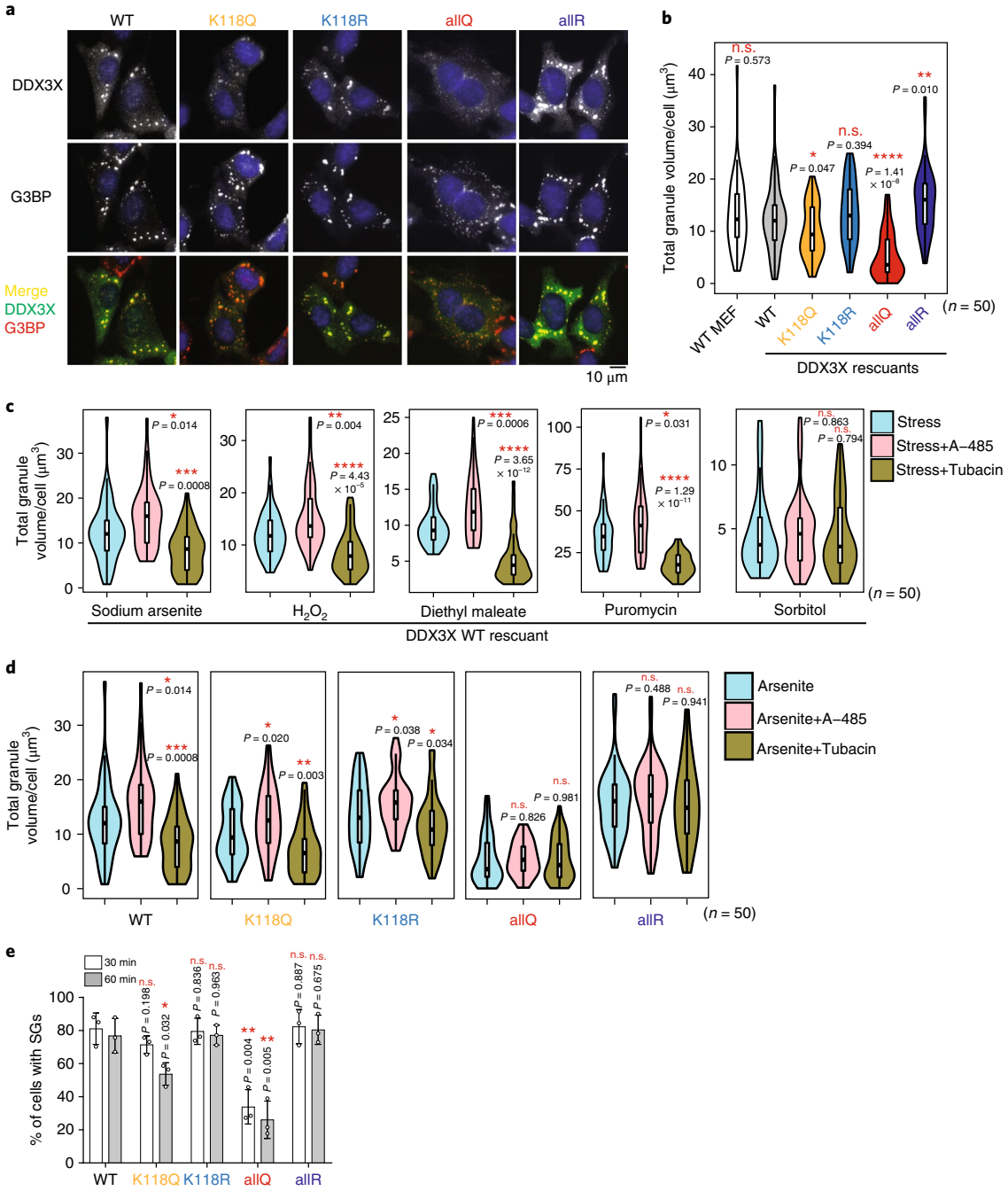


Fig. 5 | Deacetylation of DDX3X is required for normal SG size. a, Localization of DDX3X WT and mutants in DDX3X KO MEFs. FLAG-DDX3X WT, K118Q/R mutants and allQ/R mutants (all lysines in IDR1 mutated to glutamine or arginine) were stably expressed in DDX3X KO MEFs. Cells were treated with 1 mM arsenite for 1 h, and localization of FLAG-DDX3X and endogenous G3BP (SG marker) was assessed by immunofluorescence microscopy. Representative results are shown ($n = 3$ biologically independent samples). Scale bar, 10 μm. **b**, Total volume of DDX3X SGs per cell was quantified ($n = 50$ cells in total, from biologically independent experiments) from **a** and displayed in violin plot. *P* values were determined by Student's two-tailed *t*-test; $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ and $^{****}P < 0.0001$. **c**, DDX3X SG volume under different kinds of stress. After 3 h of treatment with 10 μM A-485 or 10 μM Tubacin, DDX3X WT rescuant was further treated with the indicated stresses (1 mM sodium arsenite for 1 h, 2 mM H_2O_2 for 1 h, 3 mM diethyl maleate for 1 h, 20 μg/ml puromycin for 3 h and 0.5 M sorbitol for 1 h), and localization of DDX3X was assessed by immunofluorescence microscopy. Total volume of DDX3X SGs per cell was quantified ($n = 50$ cells in total, from biologically independent experiments) and displayed in violin plot. *P* values were determined by Student's two-tailed *t*-test; $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ and $^{****}P < 0.0001$. **d**, Effects of CBP and HDAC6 inhibitors on SG formation. After 3 h of treatment with 10 μM A-485 or 10 μM Tubacin, each DDX3X rescuant was further treated with 1 mM sodium arsenite for 1 h, and localization of DDX3X was assessed by immunofluorescence microscopy. Total volume of DDX3X SGs per cell was quantified ($n = 50$ cells in total, from biologically independent experiments), and displayed in violin plot. In **b**–**d**, the thin line at the center represents the 95% confidence interval, and the bottom and top of the box represent the 25th and 75th percentile of the data, with the median value represented by a horizontal bold line. On each side of the thin line is a kernel density estimation showing the distribution shape of the data points from minima to maxima. *P* values were determined by Student's two-tailed *t*-test; $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ and $^{****}P < 0.0001$. **e**, Percentage of cells still having SGs after 30 min or 60 min post-stress recovery (mean ± s.d.; $n = 3$ biologically independent samples). *P* values were determined by Student's two-tailed *t*-test; $^{*}P < 0.05$ and $^{**}P < 0.01$.

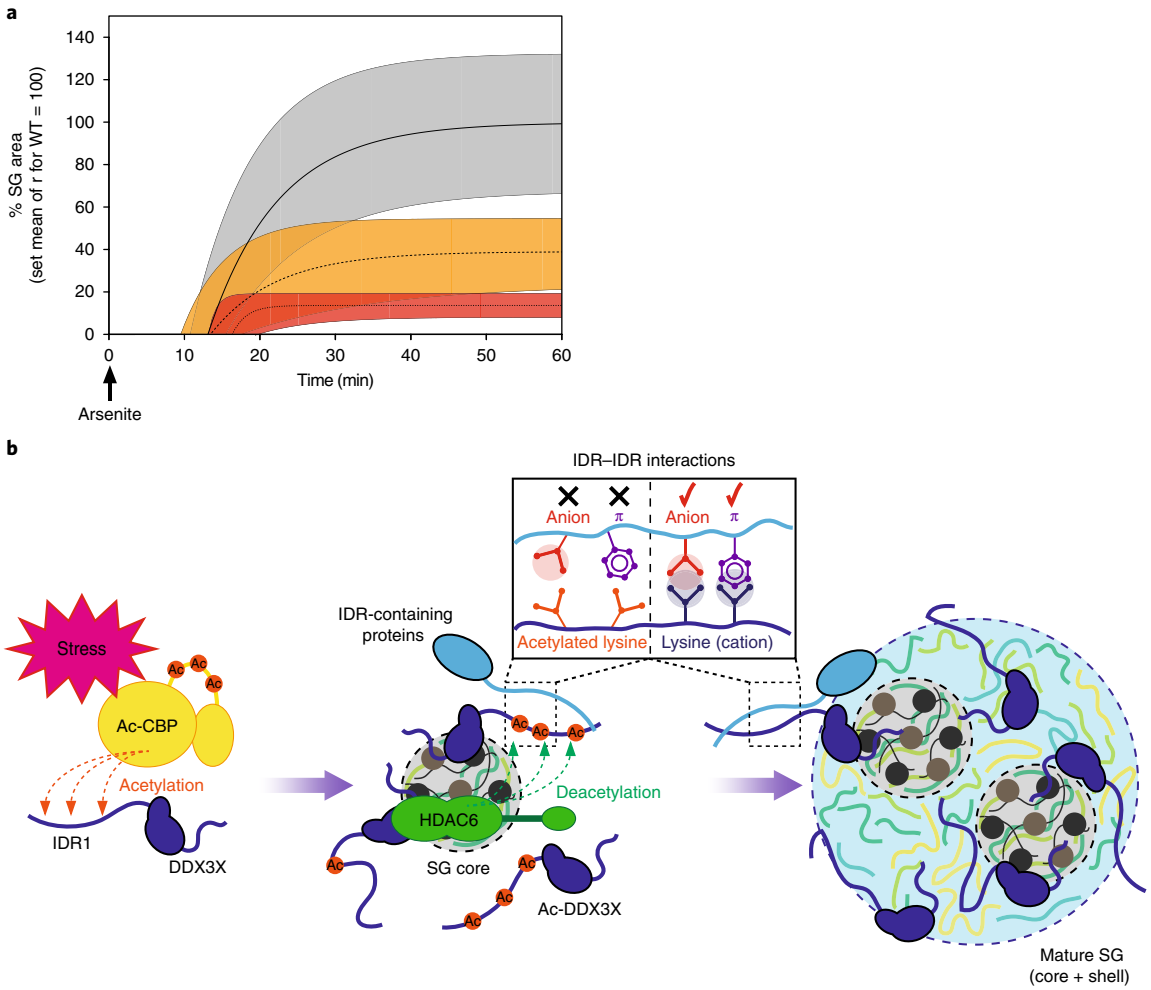


Fig. 6 | SG maturation is promoted by deacetylation of DDX3X. a, Simulation of SG growth curves for each construct with parameters in Supplementary Fig. 26a. The mean value of *r* for WT is arbitrarily set as 100, and the curves obtained with the best estimates of *T*, *k* and *r* are shown. The areas where each curve exists with three parameters (*T*, *k* and *r*) in the range of mean ± s.d. are also shown (WT: solid line ± gray zone, K118Q: dashed line ± orange zone and allQ: dotted line ± red zone). *T*, *k* and *r* represent the initiation time of SGs (*T*), their growth rate (*k*) and the ratio of SG area at steady state vs. *t* = 0. **b**, Proposed molecular model. Upon stress, CBP becomes activated and elicits acetylation of DDX3X-IDR1 among other proteins, thus impairing LLPS and leading to formation of small SGs. Though CBP does not localize to SGs, HDAC6 is recruited to SGs, where it deacetylates the IDR1. Positively charged lysine residues in the IDR1 contribute to cation–anion or cation–π interactions. This promotes LLPS of DDX3X and allows other IDR-containing proteins as interaction partners of DDX3X-IDR1 to engage in the formation of large mature SGs.

LLPS during SG maturation, and may help to further refine the models for SG assembly. Our observations also offer a novel therapeutic rationale for targeting HDACs or HATs in diverse neurodegenerative diseases.

The contribution of LLPS to biological processes in living cells is only little explored. We close by pointing out the emerging concept that transcription may also be regulated by LLPS[48–50]. Transcriptional hubs are composed of multiple components, many of which carry PTMs, and interactions between them may provide the valency necessary for LLPS. Given that HDAC6 regulates SG formation by promoting LLPS, it is tempting to speculate that other HDACs in the nucleus may also regulate the formation of transcriptional hubs by modulating LLPS. It is important to consider that acetylated lysines also provide a binding platform for acetyllysine binding "reader" domains, which are often part of transcription regulatory proteins, and may themselves promote LLPS. This could therefore be another mechanism for acetylation-mediated LLPS, distinct from our findings that acetylation impairs LLPS through neutralization of IDRs. Collectively, we provide evidence that histone deacetylases regulate LLPS in membrane-less organelle formation, and anticipate that additional biological processes regulated by acetylation-modulated LLPS will be identified in the future.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41589-018-0180-7.

Received: 14 December 2017; Accepted: 26 October 2018;
Published online: 10 December 2018

References

1. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **30**, 39–58 (2014).
2. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
3. Tompa, P. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem. Sci.* **37**, 509–516 (2012).
4. Protter, D. S. W. & Parker, R. Principles and properties of stress granules. *Trends Cell Biol.* **26**, 668–679 (2016).

5. Tyedmers, J., Mogk, A. & Bukau, B. Cellular strategies for controlling protein aggregation. *Nat. Rev. Mol. Cell Biol.* **11**, 777–788 (2010).
6. Molliex, A. et al. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 (2015).
7. Patel, A. et al. A liquid-to-solid phase transition of the ALS Protein FUS accelerated by disease mutation. *Cell* **162**, 1066–1077 (2015).
8. Brangwynne, C. P., Tompa, P. & Pappu, R. V. Polymer physics of intracellular phase transitions. *Nat. Phys.* **11**, 899–904 (2015).
9. Nott, T. J. et al. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **57**, 936–947 (2015).
10. Hofweber, M. et al. Phase separation of FUS Is suppressed by its nuclear import receptor and arginine methylation. *Cell* **173**, 706–719 e713 (2018).
11. Monahan, Z. et al. Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J.* **36**, 2951–2967 (2017).
12. Kawaguchi, Y. et al. The deacetylase HDAC6 regulates aggresome formation and cell viability in response to misfolded protein stress. *Cell* **115**, 727–738 (2003).
13. Kwon, S., Zhang, Y. & Matthias, P. The deacetylase HDAC6 is a novel critical component of stress granules involved in the stress response. *Genes Dev.* **21**, 3381–3394 (2007).
14. Zhang, Y. et al. HDAC-6 interacts with and deacetylates tubulin and microtubules in vivo. *EMBO J.* **22**, 1168–1179 (2003).
15. Hubbert, C. et al. HDAC6 is a microtubule-associated deacetylase. *Nature* **417**, 455–458 (2002).
16. Matthias, P., Yoshida, M. & Khochbin, S. HDAC6 a new cellular stress surveillance factor. *Cell Cycle* **7**, 7–10 (2008).
17. Boyault, C., Sadoul, K., Pabion, M. & Khochbin, S. HDAC6, at the crossroads between cytoskeleton and cell signaling by acetylation and ubiquitination. *Oncogene* **26**, 5468–5476 (2007).
18. Tourrière, H. et al. The RasGAP-associated endoribonuclease G3BP assembles stress granules. *J. Cell Biol.* **160**, 823–831 (2003).
19. Legros, S. et al. The HTLV-1 Tax protein inhibits formation of stress granules by interacting with histone deacetylase 6. *Oncogene* **30**, 4050–4062 (2011).
20. Schölz, C. et al. Acetylation site specificities of lysine deacetylase inhibitors in human cells. *Nat. Biotechnol.* **33**, 415–423 (2015).
21. Bannister, A. J., Miska, E. A., Görlich, D. & Kouzarides, T. Acetylation of importin-alpha nuclear import factors by CBP/p300. *Curr. Biol.* **10**, 467–470 (2000).
22. Zhang, X. et al. HDAC6 modulates cell motility by altering the acetylation level of cortactin. *Mol. Cell* **27**, 197–213 (2007).
23. Choudhary, C. et al. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834–840 (2009).
24. Zhang, L. et al. Proteomic identification and functional characterization of MYH9, Hsc70, and DNAJA1 as novel substrates of HDAC6 deacetylase activity. *Protein Cell* **6**, 42–54 (2015).
25. Shih, J. W. et al. Critical roles of RNA helicase DDX3 and its interactions with eIF4E/PABP1 in stress granule assembly and stress response. *Biochem. J.* **441**, 119–129 (2012).
26. North, B. J., Marshall, B. L., Borra, M. T., Denu, J. M. & Verdin, E. The human Sir2 ortholog, SIRT2, is an NAD^+-dependent tubulin deacetylase. *Mol. Cell* **11**, 437–444 (2003).
27. Hao, R. et al. Proteasomes activate aggresome disassembly and clearance by producing unanchored ubiquitin chains. *Mol. Cell* **51**, 819–828 (2013).
28. Grozinger, C. M., Hassig, C. A. & Schreiber, S. L. Three proteins define a class of human histone deacetylases related to yeast Hda1p. *Proc. Natl Acad. Sci. USA* **96**, 4868–4873 (1999).
29. Hai, Y. & Christianson, D. W. Histone deacetylase 6 structure and molecular basis of catalysis and inhibition. *Nat. Chem. Biol.* **12**, 741–747 (2016).
30. Miyake, Y. et al. Structural insights into HDAC6 tubulin deacetylation and its selective inhibition. *Nat. Chem. Biol.* **12**, 748–754 (2016).
31. Thompson, P. R. et al. Regulation of the p300 HAT domain via a novel activation loop. *Nat. Struct. Mol. Biol.* **11**, 308–315 (2004).
32. Floor, S. N., Condon, K. J., Sharma, D., Jankowsky, E. & Doudna, J. A. Autoinhibitory interdomain interactions and subfamily-specific extensions redefine the catalytic core of the human DEAD-box protein DDX3. *J. Biol. Chem.* **291**, 2412–2421 (2016).
33. Wang, A. et al. A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. *EMBO J.* **37**, e97452 (2018).
34. Li, X. et al. The repeat region of cortactin is intrinsically disordered in solution. *Sci. Rep.* **7**, 16696 (2017).
35. Kato, M. et al. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
36. Youn, J. Y. et al. High-density proximity mapping reveals the subcellular organization of mrna-associated granules and bodies. *Molecular cell* **69**, 517–532 e511 (2018).
37. Jain, S. et al. ATPase-modulated stress granules contain a diverse proteome and substructure. *Cell* **164**, 487–498 (2016).
38. Protter, D. S. W. et al. Intrinsically disordered regions can contribute promiscuous interactions to RNP granule assembly. *Cell Rep.* **22**, 1401–1412 (2018).
39. Wheeler, J. R., Matheny, T., Jain, S., Abrisch, R. & Parker, R. Distinct stages in stress granule assembly and disassembly. *eLife* **5**, e18413 (2016).
40. Markmiller, S. et al. Context-dependent and disease-specific diversity in protein interactions within stress granules. *Cell* **172**, 590–604 e513 (2018).
41. Riback, J. A. et al. Stress-triggered phase separation is an adaptive, evolutionarily tuned response. *Cell* **168**, 1028–1040.e1019 (2017).
42. Fanfoni, M. & Tomellini, M. The Johnson-Mehl-Avrami-Kohnogorov model: a brief review. *IlNuovo Cimento D* **20**, 1171–1182 (1998).
43. Valentin-Vega, Y. A. et al. Cancer-associated DDX3X mutations drive stress granule assembly and impair global translation. *Sci. Rep.* **6**, 25996 (2016).
44. Elbaum-Garfinkle, S. et al. The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc. Natl Acad. Sci. USA* **112**, 7189–7194 (2015).
45. Ohn, T., Kedersha, N., Hickman, T., Tisdale, S. & Anderson, P. A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly. *Nat. Cell Biol.* **10**, 1224–1231 (2008).
46. Jedrusik-Bode, M. et al. The sirtuin SIRT6 regulates stress granule formation in *C. elegans* and mammals. *J. Cell Sci.* **126**, 5166–5177 (2013).
47. Cohen, T. J. et al. An acetylation switch controls TDP-43 function and aggregation propensity. *Nat. Commun.* **6**, 5845 (2015).
48. Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
49. Chong, S. et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555 (2018).
50. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A phase separation model for transcriptional control. *Cell* **169**, 13–23 (2017).

Acknowledgements

We are grateful to T. Hyman for use of the microscope with thermal stage on short notice and for comments on the manuscript, and R. Voit (German Cancer Research Center, Heidelberg) for HAT expression vectors. We thank L. Gelman and S. Bourke for help with microscopic analysis, H. Kohler for FACS analysis, J. Seebacher and V. Iesmantavicius for interpretation of mass spectrometry data, H. Gut for help with structure predictions, M.B. Stadler for acetylome-wide IDR analysis, J. Wilbertz for help with live-cell imaging, L. Giorgetti and Y. Zhan for help with mathematical modeling, W. Filipowicz and J. Chao for critical comments on the manuscript. We thank C. Schölz for valuable suggestions. We also thank L. Wang for advice on protein purification, G. Matthias and C. Cao for their helpful technical assistance, Y. Miyake for providing us with biological materials for experiments, R. Clerc for critical comments on the manuscript, and all the Matthias laboratory members for fruitful discussions. M. Saito is supported in part by a fellowship from the Nakajima Foundation. A.W. Fritsch is supported by the ELBE postdoctoral fellows program. The Novo Nordisk Foundation Center for Protein Research is supported financially by the Novo Nordisk Foundation (Grant agreement NNF14CC0001). This work was supported by the Novartis Research Foundation.

Author contributions

M.S. and P.M. designed the project; M.S. performed all experiments and interpreted the data for the manuscript under the supervision of P.M.; M.S. and D.H. performed mass spectrometry analysis; M.S. and J.E. performed microscopy and image analysis; M.S., A.W.F. and M.K. performed temperature-dependent microscopy measurements and their image analysis; M.S. and B.T.W. analyzed HDAC6 related acetylome data under the supervision of C.C.; M.S. and P.M. wrote the manuscript and all authors contributed to the final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41589-018-0180-7.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to P.M.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

DNA constructs and antibodies. DDX3X cDNA was cloned from WT MEF cells into expression vectors (pcDNA3.1 for transient expression or pMSCV for stable expression). Expression vectors for HATs (FLAG-PCAF, HA-CBP, Myc-Tip60 and HA-p300) were kindly supplied from R. Voit (German Cancer Research Center, Heidelberg). To construct expression vectors for functional mutants of HDAC6 and DDX3X, a plasmid of interest was amplified with appropriate sets of primers, then the PCR product was self-ligated to obtain a mutated plasmid. Primary antibodies were as follows: anti-α-tubulin (DM1A; Sigma, T9026), anti-acetylated tubulin (6-11B-1) (Sigma, T7451), anti-FLAG (M2) (Sigma, F1804), anti-HDAC6 (ab56926), anti-HA (16B12) (ab130275), anti-c-Myc (9E10) (ab32), anti-phopho-eIF2α (Ser51) (D9G8) (Cell Signaling Technology (CST) #3398), anti-histone H3 (D1H2) (CST#4499), anti-acetyl-CBP (Lys1535)/p300 (Lys1499) (CST#4771), anti-eIF2α (D7D3) (CST#5324), anti-CBP (D6C5) (CST#7389), anti-HDAC6 (D21B10) (CST#7612), anti-DDX3 (D19B4) (CST#8192), anti-TIAR (D32D3) (CST#8509), anti-acetylated-lysine (CST#9441), anti-PABP1 (clone 10E10) (Millipore#04-1467), anti-acetyl-histone H3 (Millipore#06-599), anti-DDX3X (Millipore#09-860), anti-cortactin (p80/85) (4F11) (Millipore#05-180), anti-DDX3X (clone 15D1B11) (BioLegend#658602), anti-Ub (P4D1) (Santa Cruz, sc-8017), anti-[K(Ac)40]-α-tubulin (Enzo, BML-SA452-0100) and anti-G3BP (Aviva Systems Biology, ARP37713_T100). Anti-mouse-HDAC6 was developed in the Matthias laboratory in the Friedrich Miescher Institute (FMI)[51]. Secondary antibodies were as follows: Amersham ECL Mouse IgG, HRP-linked whole Ab from sheep (GE Healthcare, NA931V), Amersham ECL Rabbit IgG, HRP-linked whole Ab from donkey (GE Healthcare, NA934V), Alexa Fluor 488 goat anti-mouse IgG (H + L) secondary antibody (Invitrogen, A11001), alexa fluor 568 goat anti-mouse IgG (H + L) secondary antibody (Invitrogen, A11004), alexa fluor 488 goat anti-rabbit IgG (H + L) secondary antibody (Invitrogen, A11034) and alexa fluor 568 goat anti-rabbit IgG (H + L) secondary antibody (Invitrogen, A11036).

Establishment and characterization of DDX3X-K118Ac-specific antibody. The DDX3X-K118 acetylated peptide, CDRSGFGK(Ac)FERG (PSL Peptide Specialty Laboratories) was conjugated with mcKLH by Imject Maleimide Activated Carrier Protein Spin Kits (Thermo scientific) and the KLH conjugated peptide was used to immunize two rabbits (Pocono Rabbit Farm & Laboratory). Collected serum was passed over K118 acetylated peptide CDRSGFGK(Ac)FERG-conjugated agarose column prepared by SulfoLink Immobilization Kit for Peptides (Thermo scientific) and eluted with 0.2 M glycine-HCl, pH 2.0 and neutralized. Then, the elution was passed over K118 unacetylated peptide CDRSGFGKFERG-conjugated agarose column, and the flow through was used as a DDX3X-K118Ac-specific antibody. The specificity for K118Ac was assessed by ELISA using K118 acetylated and unacetylated peptide-coated plates (TaKaRa) and 1-Step Slow TMB-ELISA (Thermo scientific).

Cell culture. MEF cells, HEK293T cells, HeLa cells and Plat-E packaging cells were cultured in DMEM supplemented with 10% heat-inactivated FBS, penicillin and streptomycin at 37 °C and 5% CO_2.

Transfection. HEK293T cells, HeLa cells and Plat-E packaging cells were transfected with Lipofectamine 2000 (Life Technologies) or FuGENE HD (Promega), and MEF cells were transfected with 4D-Nucleofector System (Lonza), following the manufacturer's protocol.

Immunoblotting and immunoprecipitation. Cells were washed by ice-cold PBS and lysed in Triton lysis buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA, 0.1% TritonX-100 and Complete EDTA-free protease inhibitors (Roche)) for analysis. To detect protein acetylation, 0.2 μM trichostatin A and 5 mM nicotinamide were added to PBS for wash, and 10 μM trichostatin A, 10 mM nicotinamide and 50 mM sodium butyrate were added to Triton lysis buffer. To detect protein phosphorylation, Pierce Protease and Phosphatase Inhibitor Mini Tablets (Thermo Scientific) were added to Triton lysis buffer instead of Complete EDTA-free protease inhibitors. 500 μg protein lysates were incubated with the specific antibody overnight at 4 °C, and immunoprecipitated with Dynabeads Protein G (Invitrogen) for 1 h at 4 °C. Samples were boiled for 10 min in SDS–PAGE sample buffer, and separated with 4–12% NuPAGE gels (Invitrogen). Proteins were transferred onto PVDF membranes (Immobilon-P, Millipore), probed with specific primary antibody overnight and secondary antibody for 1 h under 5% non-fat dry milk in TBS or 5% BSA, 0.1% Tween20 in TBS blocking conditions and detected with Amersham ECL western blotting reagent (GE Healthcare). For inhibition of HDACs in cells, the concentration of each HDAC inhibitor was as described in ref. [20].

Identification of the DDX3X acetylation sites. The protein pellets were reduced with 10 mM TCEP, alkylated with 20 mM iodoacetamide and cleaved first with 0.1 μg Lys-C (Wako) for 6 h and then with 0.1 μg porcine sequencing grade trypsin (Promega) overnight. The extracted peptides were analyzed by capillary liquid chromatography tandem mass spectrometry with an EASY-nLC 1000 using the two-column set up (Thermo Scientific). The peptides were loaded in 0.1% formic acid, 2% acetonitrile in H_2O onto a peptide trap (Acclaim PepMap 100, 75 μm × 2 cm, C18, 3 μm, 100 Å) at a constant pressure of 600 bar. Then they were separated, at a flow rate of 150 nl/min with a linear gradient of 2–6% buffer B in buffer A in 3 min followed by a linear increase from 6 to 22% in 40 min, 22–28% in 9 min, 28–36% in 8 min, 36–80% in 1 min and 12 min at 80% buffer B in buffer A (buffer A: 0.1% formic acid, buffer B: 0.1% formic acid in acetonitrile) on a 50 μm × 15 cm ES801 C18, 2 μm, 100 Å column mounted on a DPV ion source (New Objective) connected to a Orbitrap Fusion (Thermo Scientific). The data were acquired using 120,000 resolution for the peptide measurements in the Orbitrap and a top T (3 s) method with HCD fragmentation for each precursor and fragment measurement in the linear ion trap. MASCOT 2.5 was used in the Decoy mode to search the Swiss-Prot mouse version 2015_01 including common contaminants. The enzyme specificity was set to trypsin, allowing for up to three incomplete cleavage sites. Carbamidomethylation of cysteine (+ 57.0245 Da) was set as a fixed modification, oxidation of methionine (+ 15.9949 Da) and acetylation of lysines (+ 42.0106 Da) were set as variable modifications. Parent ion mass tolerance was set to 10 p.p.m. and fragment ion mass tolerance to 0.6 Da. The results were validated with the program Scaffold (version: 4.4) and Scaffold PTM (version: 2.2) (Proteome Software, Portland, USA). Peptide identifications were accepted if they could be established at greater than 0.1% probability as calculated in Scaffold and acetylation sites were accepted if they had a greater than 80% site probability as calculated with Scaffold PTM.

Interaction proteomics. FLAG-DDX3X was immunoprecipitated with Anti-FLAG M2 Magnetic Beads (Sigma) from cellular lysate of each 293T sample expressing either DDX3X WT or mutants. Cellular lysate of parental 293T cells was used as a negative control. Precipitants were digested on beads by Lys-C (Wako) in digestion buffer (3 M GuHCl, 20 mM EPPS, pH 8.5, 10 mM CAA and 5 mM TCEP), further digested by trypsin (Promega), and analyzed by LC–MS/MS, essentially as described[52]. In short, the peptides were separated with an EASY-nLC 1000 on a 50 μm × 15 cm ES801 C18, 2 μm, 100 Å column (Thermo Scientific) mounted on a DPV ion source (New Objective). They were measured with a Orbitrap Fusion (Thermo Scientific) using a top T (3 s) method as recommended by the manufacturer (Thermo Scientific). Andromeda implemented in MaxQuant (version: 1.5.3.8)[53] was used to search the mouse subset of the UniProt (version: 2017_04) combined with the contaminant database from MaxQuant and label-free quantification (LFQ[54]) was used with a protein and peptide FDR of 0.01. Statistical analysis was done in Perseus (version: 1.5.2.6)[55,56]. Results were filtered to remove reverse hits, contaminants and peptides found in only one sample. Missing values were imputed and potential interactors were determined using Student's two-tailed *t*-test and visualized by a volcano plot. Significance lines corresponding to a FDR of 0.03 and a S0 of 1.5 are shown in the corresponding figures. Results were exported from Perseus and visualized using statistical computing language R.

Biotinylated isoxazole (b-isox)-mediated precipitation. Cells were washed by ice-cold PBS and lysed in EE buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM EDTA, 2.5 mM EGTA, 0.1% NP-40, 10% Glycerol, 1 μM DTT and Complete EDTA-free protease inhibitors (Roche)) To detect DDX3X acetylation, 10 μM Tubacin was added to EE buffer. Biotinylated isoxazole (Sigma) in DMSO was added to cell lysates at 100 μM final concentration. The reaction solutions were incubated for 4 h at 4 °C and then centrifuged at 13,200 r.p.m. for 10 min. Precipitates were washed five times in EE buffer before SDS solubilization, following analyzed by immunoblotting as mentioned above.

Immunofluorescence microscopy and live-cell imaging. Stress granules were induced by the following conditions: 1 mM sodium arsenite (Millipore) for 15 min–1 h, 2 mM H_2O_2 (Sigma) for 1 h, 3 mM diethyl maleate (Sigma) for 1 h, 20 μM CCCP (Sigma) in glucose-free DMEM (Gibco) with 10% FBS for 1 h, 40 μM clotrimazole (Sigma) in glucose-free DMEM with 10% FBS for 3 h, 20 μg/ml puromycin (Sigma) for 3 h, 10 μM thapsigargin (Invitrogen) for 1 h, 10 μM MG132 (Sigma) for 3 h, heat shock at 43 °C for 1 h and 0.5 M sorbitol (Sigma) for 1 h. Catalytic activities of HDAC6 or CBP were inhibited by treatment of 10 μM Tubacin (Sigma) or 10 μM A-485 (Lucerna-Chem) for 3 h, respectively. For knockdown of CBP, ON-TARGETplus SMARTpool siRNA mouse CBP (Dharmacon) were delivered via nucleofection (2 μM). Cells on a coverglass were washed with ice-cold PBS, then fixed with ice-cold methanol or 4% paraformaldehyde. After permeabilization with 0.5% TritonX-100 or 0.005% digitonin in PBS for 10 min, the cells were incubated with specific primary and secondary antibodies in 10% goat serum (Sigma) blocking buffer, then mounted with ProLong Gold Antifade Reagent with DAPI (Cell Signaling Technology #8961). Images were captured by Axioimager Z1 microscope (Zeiss). SGs in images were defined by a plugin of Fiji (SG counter). Volume of SG $V = 4\pi R^3/3$, where R is SG radius, was calculated for each SG, and total volume of SGs in a cell was determined. In each violin plot, the thin line at the center represents the 95% confidence interval, and the bottom and top of the box represent the 25th and 75th percentile of the data, with the median value represented by a horizontal bold line. On each side of the thin line is a kernel density estimation showing the distribution shape of the data points from minima to maxima. For live-cell imaging, WT or DDX3X KO MEF cells were transfected with expression vectors encoding mCherry-DDX3X WT or mutants, cultured for 1 d on a glass-bottom dish 35 mm

(Ibidi), and images were captured with a spinning-disk confocal microscope (Olympus IX81 + Yokogawa CSU-X1 scan head) with a 561 nm laser for mCherry signals.

Mathematical modeling of SG growth. Live-cell image stacks, captured as mentioned above, were subjected to maximum intensity projection, and SGs were defined by SG counter. The time point when cells are treated with arsenite was set as $t=0$, and the initiation time of SGs (T) was defined as the time point when SG is observed for the first time. The obtained total SG area is regressed on Avrami equation:

$$A(t) = A_F(1 - e^{-k(t-T)})$$

Two parameters, A_F and k, were estimated from this fit. To simulate kinetics of each DDX3X construct, the effect of size scaling was adjusted: the initial area A_0 was measured at $t=0$ for each cell, and the scaling ratio r ($=A_F/A_0$) was obtained. The best estimate and associated error (s.d.) for k and r were calculated from the fit ($n=5$ for each construct). To visualize the data, the mean value of r for WT is arbitrarily set as 100, and the curves obtained with the best estimates of T, k and r were shown. The areas where each curve exists with three parameters: T, k and r in the range of mean ± s.d. were also shown.

Establishment of KO cell lines using CRISPR–Cas9 genome editing. HDAC6 KO HEK293T cells were established by CRISPR–Cas9 genome editing. The guide sequence targeting human HDAC6 was designed by the CRISPR design tool at http://www.genome-engineering.org/crispr/ and cloned into pX330-Cas9-T2A-mCherry vector (Addgene). The following 20 nucleotides were inserted:

5′-GGTGGAATCCTGGCCGGTTG-3′

The pX330 vector was transfected into HEK293T cells as mentioned above. Two days after transfection, the mCherry positive cells were collected by FACS. Then, single cell clones were sorted again week after and expanded. Screening for HDAC6 knockout was done by immunoblotting. Genomic DNA was purified from the HDAC6 KO clone and the region surrounding PAM of the sgRNA was cloned into pGEM-T Easy vector (Promega) for sequencing after amplification with a pair of primers:

5′-GGCAGAGAGGTGGGGTCCTC-3′

5′-CTCATCAGGGCGGGTCTAGG-3′

To determine the indels of individual alleles, the amplicons from 20 bacterial colonies were sequenced.

Establishment of DDX3X KO MEFs followed the same procedure. The inserted nucleotides and a pair of primers for genotyping are followings:

5′-tggcagtggaaaatgcgct -3′ (for guide RNA)

5′-GAGGAGGGCACACGTCTTAC-3′ (for genotyping)

5′-AACTTAAAGAGCTGCGCCAC-3′ (for genotyping)

Protein expression and purification. The DNA fragments of mouse DDX3X-IDR1 (a.a. 2–167) and DDX3X-full-length (a.a. 2–662) were inserted into the BamHI sites of the pET28 vector. All proteins were expressed in *E. coli* BL21 (DE3) cells with 1.0 mM IPTG at 16 °C for overnight. Cells were lysed with a lysis buffer containing 20 mM Tris, pH 7.5, 500 mM NaCl, 2 mM DTT, Complete EDTA-free protease inhibitors (Roche) for 30 min on ice, followed by sonication. The lysate was centrifuged at 15.000 r.p.m. for 30 min. The supernatant was incubated with Ni-NTA resin (QIAGEN) for 1 h at 4 °C. The Ni-NTA resin was packed in a column and washed with a buffer containing 20 mM Tris, pH 7.5, 500 mM NaCl, 20 mM imidazole, 2 mM DTT. The bound proteins were eluted with a buffer containing 20 mM Tris, pH 7.5, 500 mM NaCl, 500 mM imidazole, 2 mM DTT. The purified His-SUMO-tagged proteins were concentrated with Amicon Ultra centrifugal filters (Millipore), and digested by SUMO protease for overnight at 4 °C. Then, the digested samples were passed over Ni-NTA resin and the flow through was collected and concentrated. They were subjected to size exclusion chromatography on a Superdex 200 Increase 10/300 GL column (GE Healthcare) equilibrated in the Gel Filtration buffer, 20 mM Tris, pH 7.5, 250 mM NaCl, and 2 mM DTT. The fractions were concentrated to 20 mg/ml and stored at −80 °C.

For full-length protein, 1 M NaCl buffers were used through the whole purification process. The purities of the purified proteins were tested on SDS–PAGE gels, and the concentrations were determined by absorbance at UV at 280 nm.

In vitro droplet formation assay. For the phase separation of purified DDX3X, experiment was performed in 20 mM Tris, pH 7.5, 2 mM DTT. The concentrations of DDX3X and NaCl are 0 – 200 μM and 0 – 250 mM, respectively as indicated in the figure legends. As a molecular crowder, PEG4000 (Sigma) was added to samples at 10% (w/v). For acetylation of DDX3X-IDR1 in vitro, 40 μg of the IDR1 was incubated with 2.5 μg of human recombinant CBP catalytic domain (Enzo) and 10 mM acetyl-CoA (Sigma) for 3 h at 37 °C. For deacetylation of DDX3X-IDR1 in vitro, *D. rerio* HDAC6 CD1-CD2 (a.a. 40-831) was expressed in Sf9 insect cells, purified based on the method previously described[30] (buffer and salt were adjusted to the droplet formation assay), and added into DDX3X-IDR1 acetylation reaction mixture. Alexa Fluor 488 Protein Labeling Kit (Molecular Probes) was used for fluorescent labeling of the IDRs, followed by dialysis against Gel Filtration buffer using Slide-A-Lyzer MINI Dialysis Devices (Thermo scientific). The droplets were observed using an Axioimager Z1 microscope (Zeiss) and Spinning-Disk Confocal microscope, and turbidity of the samples (OD_{600}) is quantified with NanoDrop (Thermo scientific). For temperature-dependent droplet dissolution and condensation experiments, samples were mounted inside small compartments sealed with a pegylated coverslip on a custom-built temperature-controlled stage[57]. For imaging, the heating/cooling system was mounted on a Nikon TiE stand equipped with a Perfect Focus System. Image stacks were captured every 10 s using a 40x air objective (0.95NA) combined with a spinning-disk confocal unit CSU-X1 (Yokogawa) and an EMCCD camera IXON + (Andor). Image data was analyzed with MATLAB. In brief, image stacks were subjected to maximum intensity projection, the background was subtracted, and they were binarized after further filtering. In the binarized image-series, particle number was analyzed.

Computational prediction of disordered regions. Computational prediction of disordered regions was done with the PONDR VSL2 and VL3-BA programs. As queries for the prediction, protein sequences from the UniProt database released in March 2017 were used. Missing protein entries in this version were cited from the one released in February 2012.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Published research reagents from the FMI are shared with the academic community under a Material Transfer Agreement (MTA) having terms and conditions corresponding to those of the UBMTA (Uniform Biological Material Transfer Agreement).

References

51. Banerjee, I. et al. Influenza A virus uses the aggresome processing machinery for host cell entry. *Science* **346**, 473–477 (2014).
52. Ostapcuk, V. et al. Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. *Nature* **557**, 739–743 (2018).
53. Cox, J. et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).
54. Cox, J. et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513–2526 (2014).
55. Hubner, N. C. et al. Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J. Cell Biol.* **189**, 739–754 (2010).
56. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
57. Mittasch, M. et al. Non-invasive perturbations of intracellular flow reveal physical principles of cell organization. *Nat. Cell Biol.* **20**, 344–351 (2018).

nature research

Corresponding author(s): Patrick Matthias

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	ZEN blue 2012 and VisiView (version: 4.1.0.4)
Data analysis	R (v: 3.4.4), Fiji (v: 2013 Jul 15), MATLAB (v: 9.3), Cytoscape (v: 3.3.0), Scaffold (v: 4.4-4.6) , ScaffoldPTM (v: 2.2), MaxQuant (v: 1.5.3.8) and Perseus (v: 1.5.2.6)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Published research reagents from the FMI are shared with the academic community under a Material Transfer Agreement (MTA) having terms and conditions corresponding to those of the UBMTA (Uniform Biological Material Transfer Agreement)

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Experiments were done with independent biological samples (usually n≥3) and highly comparable results were obtained. This sample size is widely accepted as sufficient when using independent biological samples.
Data exclusions	No data was excluded.
Replication	All attempts at replication were successful in ≥3 experiments with independent biological samples. The methods and reagents used are described in detail, so that replication of the experiments by other scientists should be facilitated.
Randomization	Samples were organized by experimental variable, measurements were done in an automated manner (e.g. by the microscope) and the results are reported in completion.
Blinding	Not applicable. The setup was the same for all samples tested in the experiment, and the analysis of obtained data was performed in parallel in an automated manner (e.g. by the microscope), with identical parameters for all measurements.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Primary antibodies were as follows: Anti-α-tubulin (DM1A) (Sigma, T9026, lot:049K4767), Anti-Acetylated Tubulin (6-11B-1) (Sigma, T7451, lot: 103M4772V), Anti-FLAG (M2) (Sigma, F1804, lot: SLBV9325), Anti-HDAC6 (ab56926, lot: GR211012-2), Anti-HA (16B12) (ab130275, lot: GR250145-3), Anti-c-Myc (9E10) (ab32, lot:931293), Anti-Phopho-eIF2α (Ser51) (D9G8) (Cell Signaling Technology (CST) #3398, lot:6), Anti-Histone H3 (D1H2) (CST#4499, lot:1), Anti-Acetyl-CBP (Lys1535)/p300 (Lys1499) (CST#4771, lot:5), Anti-eIF2α (D7D3) (CST#5324, lot:3), Anti-CBP (D6C5) (CST#7389, lot:3), Anti-HDAC6 (D21B10) (CST#7612, lot:2), Anti-DDX3 (D19B4) (CST#8192, lot:1), Anti-TIAR (D32D3) (CST#8509, lot:1), Anti-Acetylated-Lysine (CST#9441, lot:11), Anti-PABP1 (clone 10E10) (Millipore#04-1467, lot:2349625), Anti-acetyl-Histone H3 (Millipore#06-599, lot:2370129), Anti-DDX3X (Millipore#09-860, lot:2714517), Anti-Cortactin(p80/85) (4F11) (Millipore#05-180, lot:2210407), Anti-DDX3X (clone 15D1B11) (BioLegend#658602, lot:B177028), Anti-Ub (P4D1) (Santa Cruz, sc-8017, lot:G2303), Anti-[K(Ac)40]-α-tubulin (Enzo, BML-SA452-0100, lot:01051550) and Anti-G3BP (Aviva Systems Biology, ARP37713_T100, lot:QC6154-42482). Anti-mouse-HDAC6 was developed in Matthias laboratory, FMI (Rabbit polyclonal, Banerjee I. et al., Science, 346, 2014). Anti-Acetylated-DDX3X (K118) was also developed in Matthias laboratory. DDX3X-K118 acetylated peptide (PSL Peptide Speciality Laboratories) conjugated with mcKLH was used to immunize two rabbits (Pocono Rabbit Farm & Laboratory), and the fraction of interest was affinity purified from the collected serum. The specificity of the antibody for K118Ac was assessed by ELISA. Secondary antibodies were as follows: Amersham ECL Mouse IgG, HRP-linked whole Ab from sheep (GE Healthcare, NA931V, lot:12227046), Amersham ECL Rabbit IgG, HRP-linked whole Ab from donkey (GE Healthcare, NA934V, lot:13601187), Alexa Fluor 488 Goat anti-Mouse IgG (H+L) Secondary Antibody (invitrogen, A11001, lot:1858182), Alexa Fluor 568 Goat anti-Mouse IgG (H+L) Secondary Antibody (invitrogen, A11004, lot:1419715), Alexa Fluor 488 Goat anti-Rabbit IgG (H+L) Secondary Antibody (invitrogen, A11034, lot:1910795D) and Alexa Fluor 568 Goat anti-Rabbit IgG (H+L) Secondary Antibody (invitrogen, A11036, lot:1704462).
Validation	The data are provided in the manuscript. Anti-α-tubulin (DM1A) (Sigma, T9026; WB for human and mouse proteins (Dilution

1:5000)), Anti-Acetylated Tubulin (6-11B-1) (Sigma, T7451; WB for human (1:5000)), Anti-FLAG (M2) (Sigma, F1804; WB (1:2000), IP (1:100)), Anti-HDAC6 (ab56926; IF for human (1:20)), Anti-HA (16B12) (ab130275; WB (1:5000)), Anti-c-Myc (9E10) (ab32; WB (1:1000)), Anti-Phopho-eIF2α (Ser51) (D9G8) (Cell Signaling Technology (CST) #3398; WB for human and mouse (1:1000)), Anti-Histone H3 (D1H2) (CST#4499; WB for human (1:1000)), Anti-Acetyl-CBP (Lys1535)/p300 (Lys1499) (CST#4771; WB for mouse (1:1000)), Anti-eIF2α (D7D3) (CST#5324; WB for human and mouse (1:5000)), Anti-CBP (D6C5) (CST#7389; WB for mouse (1:1000), IP for mouse (1:100)), Anti-HDAC6 (D21B10) (CST#7612; WB for human (1:1000)), Anti-DDX3 (D19B4) (CST#8192; WB for human and mouse (1:2000), IP for human (1:100)), Anti-TIAR (D32D3) (CST#8509; WB for human (1:1000)), Anti-Acetylated-Lysine (CST#9441; WB for human and mouse (1:1000)), Anti-PABP1 (clone 10E10) (Millipore#04-1467; IF for human (1:100)), Anti-acetyl-Histone H3 (Millipore#06-599; WB for human (1:1000)), Anti-DDX3X (Millipore#09-860; IF for human (1:100)), Anti-Cortactin(p80/85) (4F11) (Millipore#05-180; WB for human (1:1000), IP for human (1:100)), Anti-DDX3X (clone 15D1B11) (BioLegend#658602; IF for mouse (1:100)), Anti-Ub (P4D1) (Santa Cruz, sc-8017; WB for human (1:500)), Anti-[K(Ac)40]-α-tubulin (Enzo, BML-SA452-0100; WB for human (1:5000)), Anti-G3BP (Aviva Systems Biology, ARP37713_T100; WB for mouse (1:1000), IF for human and mouse (1:100)), Anti-mouse-HDAC6 (Matthias laboratory; WB for mouse (1:1000)) and Anti-Acetylated-DDX3X (K118) (Matthias laboratory; WB for human and mouse (1:1000), IF for mouse (1:20), ELISA for synthesized peptide (1:1000)). Amersham ECL Mouse IgG, HRP-linked whole Ab from sheep (GE Healthcare, NA931V; WB (1:2000)), Amersham ECL Rabbit IgG, HRP-linked whole Ab from donkey (GE Healthcare, NA934V; WB (1:2000)), Alexa Fluor 488 Goat anti-Mouse IgG (H+L) Secondary Antibody (invitrogen, A11001; IF (1:500)), Alexa Fluor 568 Goat anti-Mouse IgG (H+L) Secondary Antibody (invitrogen, A11004; IF (1:500)), Alexa Fluor 488 Goat anti-Rabbit IgG (H+L) Secondary Antibody (invitrogen, A11034; IF (1:500)) and Alexa Fluor 568 Goat anti-Rabbit IgG (H+L) Secondary Antibody (invitrogen, A11036; IF (1:500)).

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	MEFs were generated in the Matthias laboratory. HEK293T and HeLa cells were from Matthias lab aliquots (originally obtained from ATCC). Plat-E cells, originally from the University of Tokyo, were obtained from the lab of S. Gaubatz (University of Würzburg). Sf9 cells were obtained from the lab of N. Thomä (FMI).
Authentication	Cells initially obtained from ATCC had been authenticated by the supplier (morphology, karyotyping, PCR).
Mycoplasma contamination	No mycoplasma contamination was confirmed by the authors.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

nature chemical biology

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41589-018-0180-7

In the format provided by the authors and unedited.

Acetylation of intrinsically disordered regions regulates phase separation

Makoto Saito[1,2], Daniel Hess[1], Jan Eglinger[1], Anatol W. Fritsch[3,4], Moritz Kreysing[3,4], Brian T. Weinert[5], Chunaram Choudhary[5] and Patrick Matthias[1,2*]

[1]Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland. [2]Faculty of Sciences, University of Basel, Basel, Switzerland. [3]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany. [4]Center for Systems Biology Dresden, Dresden, Germany. [5]Department of Proteomics, The Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. *e-mail: patrick.matthias@fmi.ch

SUPPLEMENTARY INFORMATION

Acetylation of intrinsically disordered regions regulates phase separation

Makoto Saito[1, 2], Daniel Hess[1], Jan Eglinger[1], Anatol W. Fritsch[3, 4], Moritz Kreysing[3, 4], Brian T. Weinert[5], Chunaram Choudhary[5] & Patrick Matthias[1, 2*]

Affiliations:

[1] Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland.

[2] Faculty of Sciences, University of Basel, 4031 Basel, Switzerland.

[3] Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany.

[4] Center for Systems Biology Dresden, 01307 Dresden, Germany.

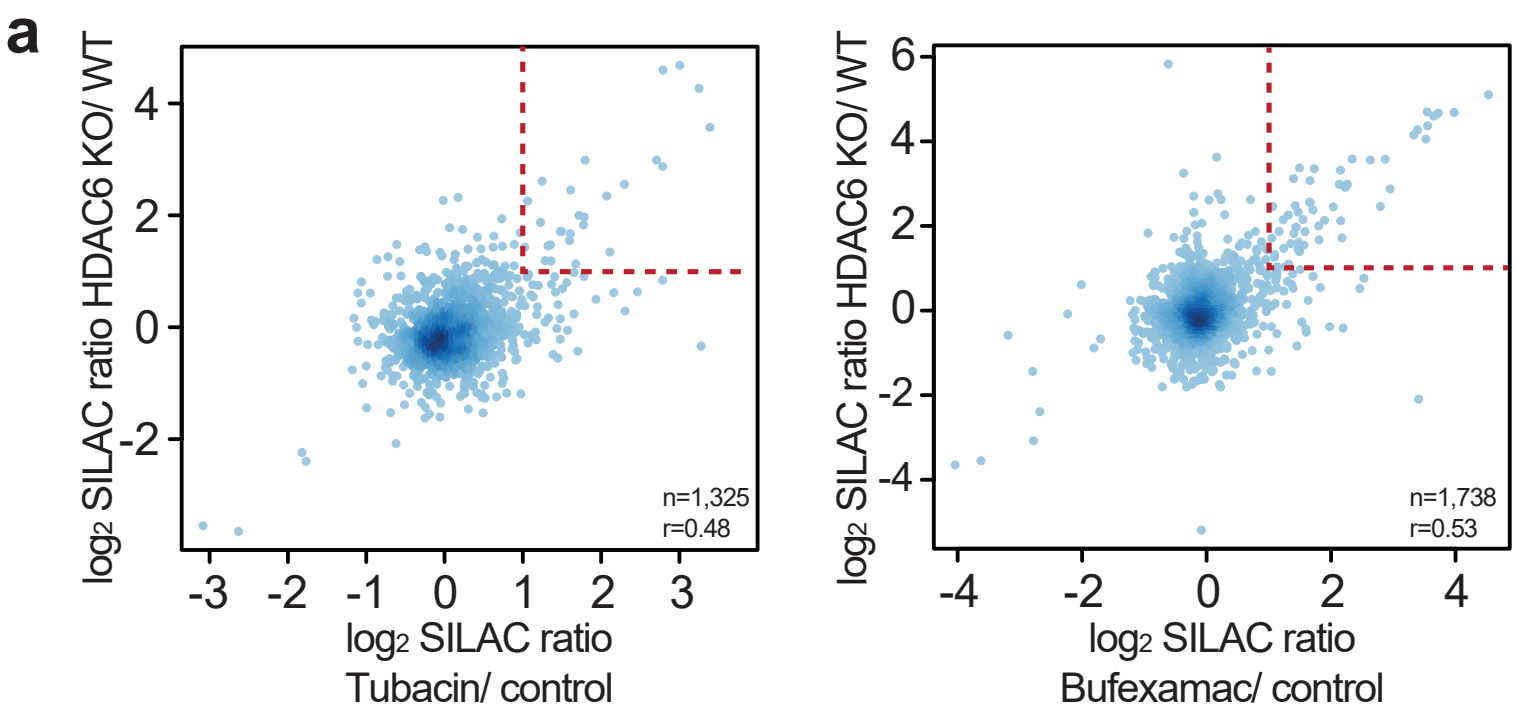
[5] Department of Proteomics, The Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen, Denmark.

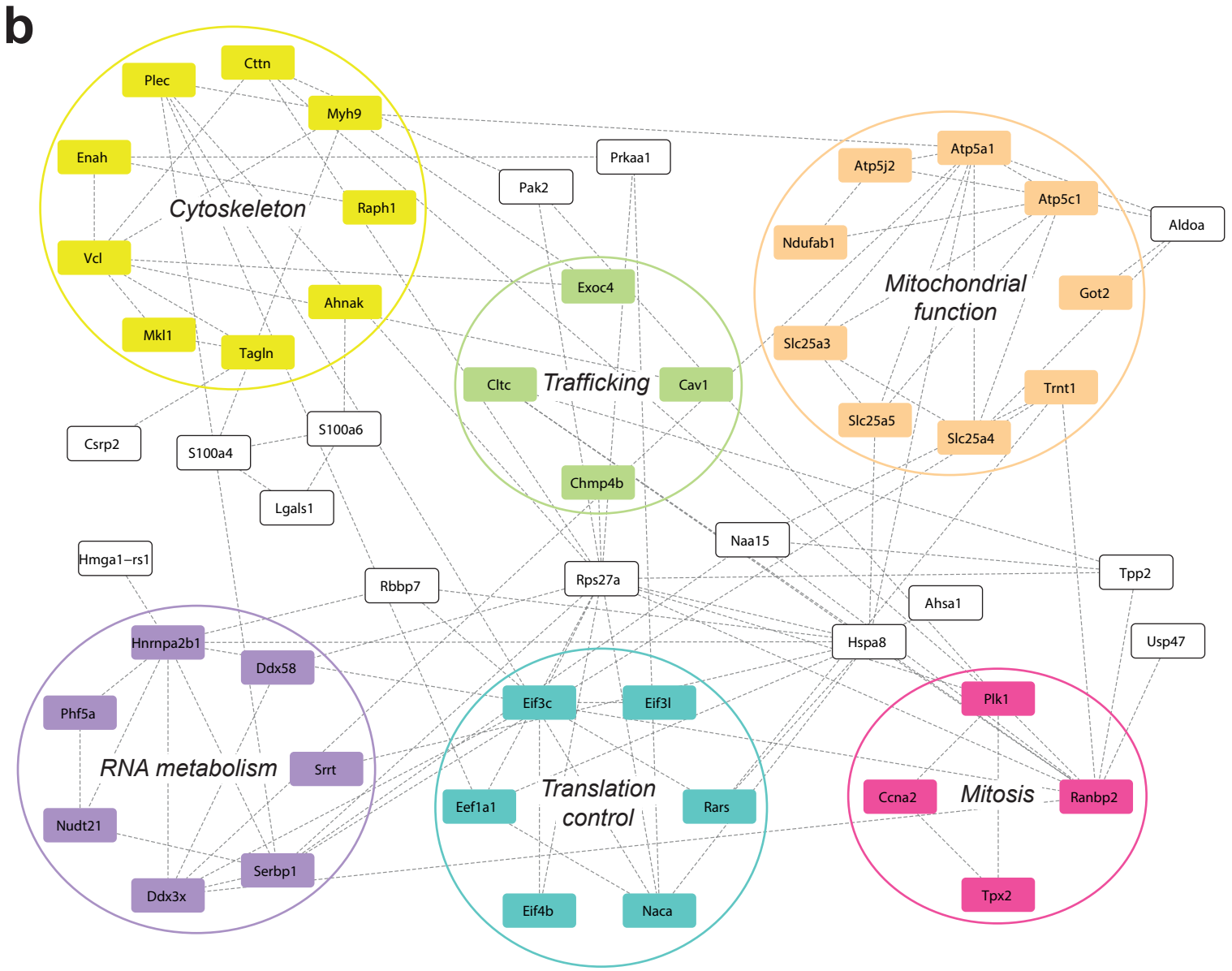
***Correspondence and Lead Contact:**

Patrick Matthias, Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland, Phone: 41-61-697-6661, FAX: 41-61-697-3976, E-mail: patrick.matthias@fmi.ch

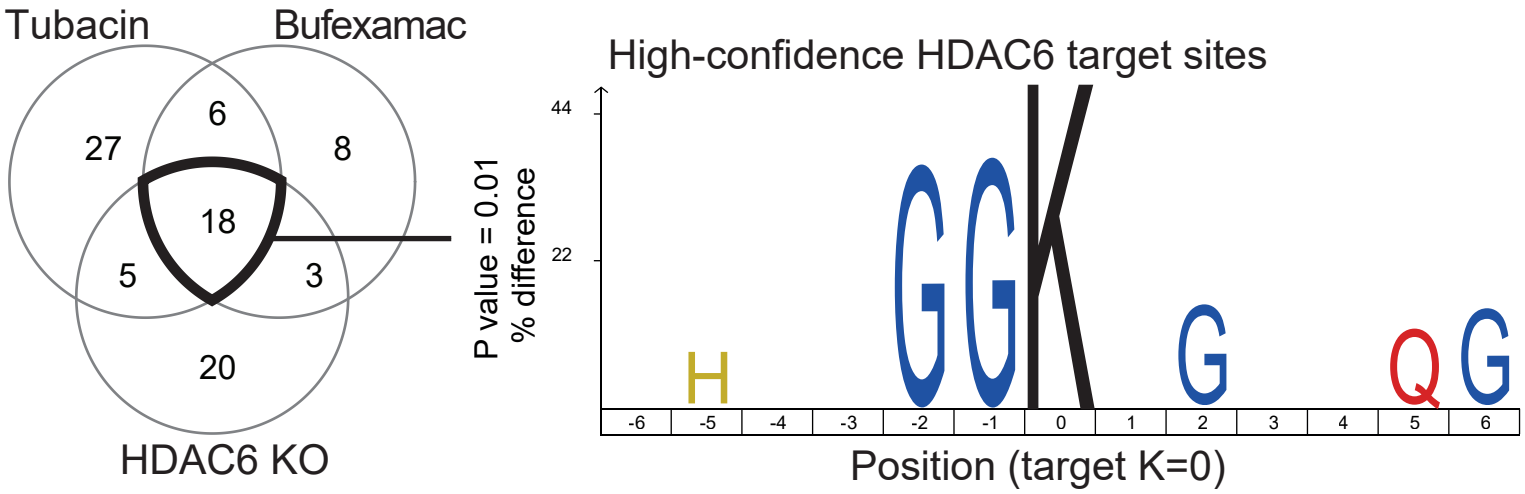
This PDF file includes:

Supplementary Fig. 1_1





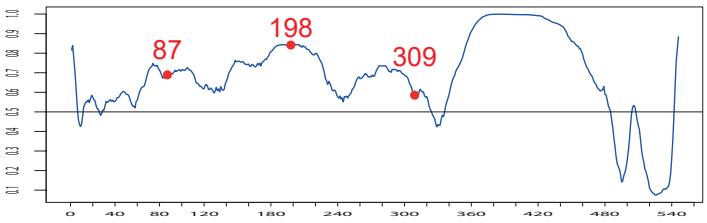
Supplementary Fig. 1_2



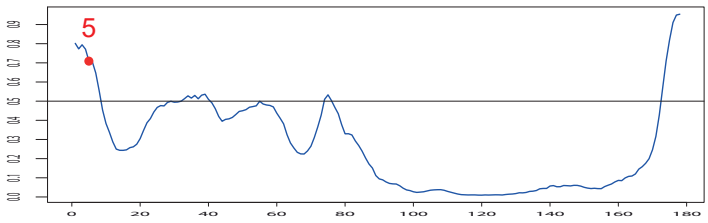
Supplementary Figure 1. Characterization of the HDAC6-dependent acetylome (related to Figure 1).

(**a**) HDAC6-dependent acetylome. The scatter plots show the correlation between acetylated sites quantified in HDAC6-specific inhibitor-treated MEFs vs acetylated sites quantified in HDAC6 KO MEFs. X-axis: SILAC ratio of Tubacin-treated MEFs vs control-treated MEFs (left panel) or SILAC ratio of Bufexamac-treated MEFs vs control-treated MEFs (right panel). Y-axis, both panels: SILAC ratio of HDAC6 KO MEFs vs WT MEFs. In each plot, the red dotted line shows the 2-fold increase threshold. Correlation was determined with Pearson correlation coefficient. (**b**) Proteins were assigned to cellular function based on gene ontology annotations using the STRING database [58]. The presented analysis was performed with the 87 peptides that showed a more than 2-fold increase in acetylation upon chemical inhibition or genetic ablation of HDAC6. The protein interaction network obtained under medium confidence (interaction score >= 0.400) is shown. (**c**) Left: Venn diagram showing the overlap between the different conditions tested: Tubacin, Bufexamac, or HDAC6 KO. The region of overlap between all 3 conditions contains 18 peptides and corresponds to the high-confidence sites. Right: Definition of an HDAC6 target motif based on analysis of the 18 high-confidence sites. The analysis was done with the iceLogo software [59]. K represents the acetylated lysine site at position 0; it has been inserted in the figure for visualization only (there is no score on the Y-axis at position 0, because all analysed peptides have K).

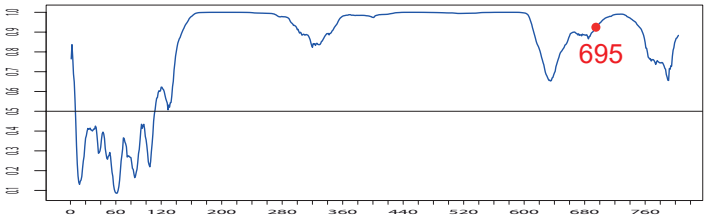
Supplementary Fig. 2



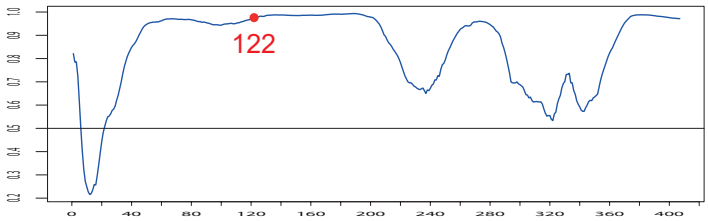
Cttn



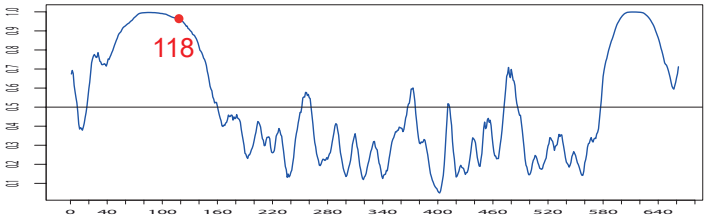
Cav1



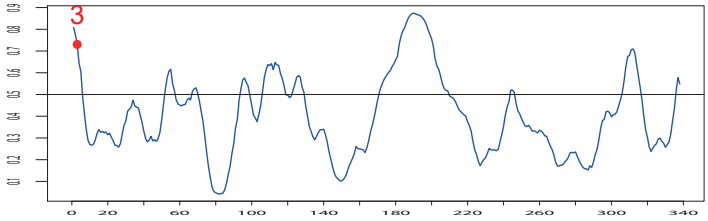
Enah



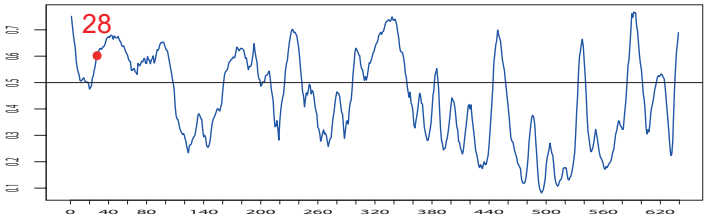
Serbp1



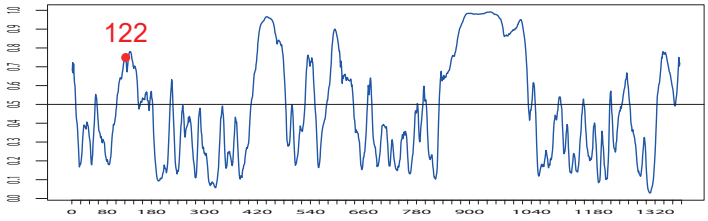
Ddx3x



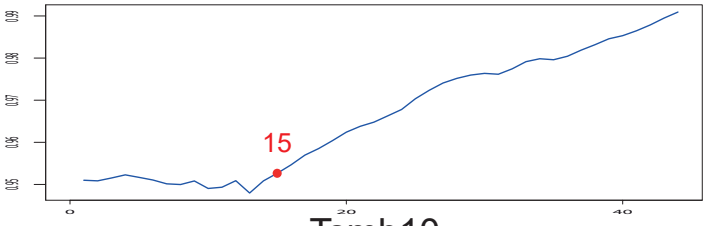
Ahsa1



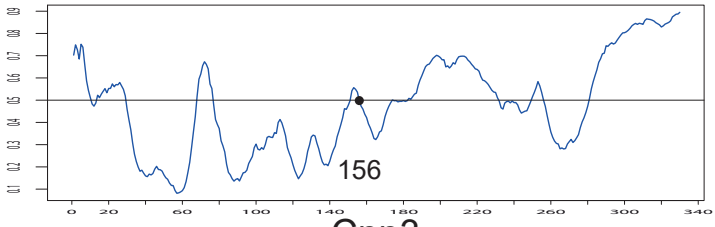
Pex5



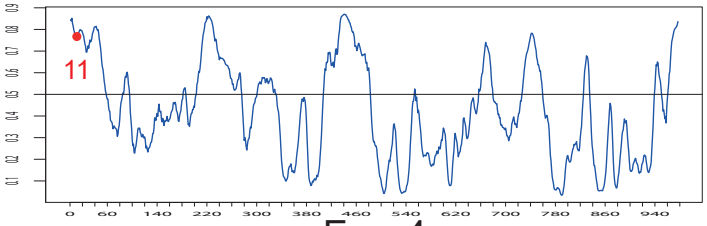
Usp47



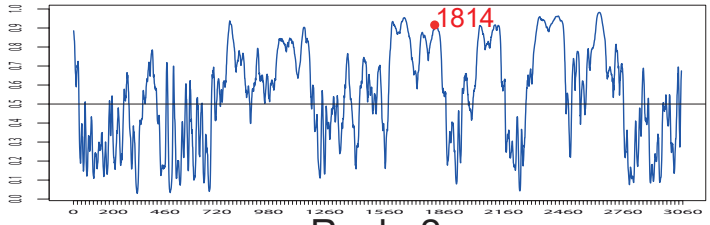
Tsmb10



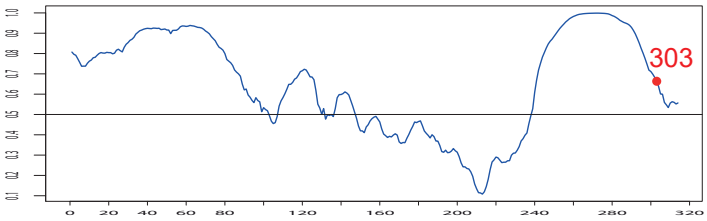
Cnn3



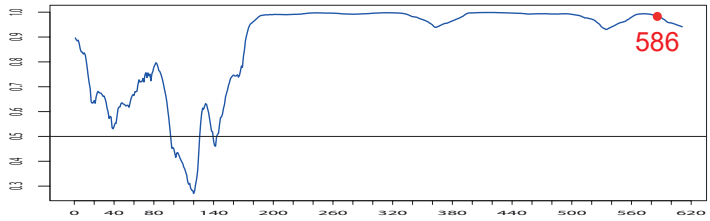
Exoc4



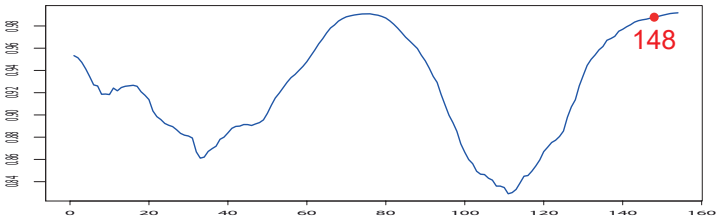
Ranbp2



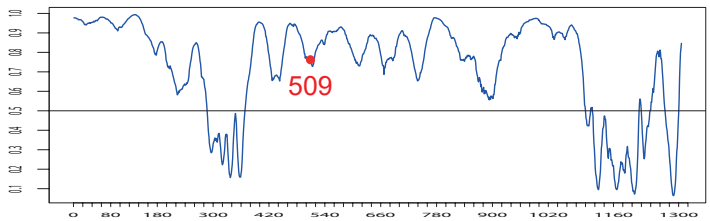
Nat6



Eif4b



Hn1

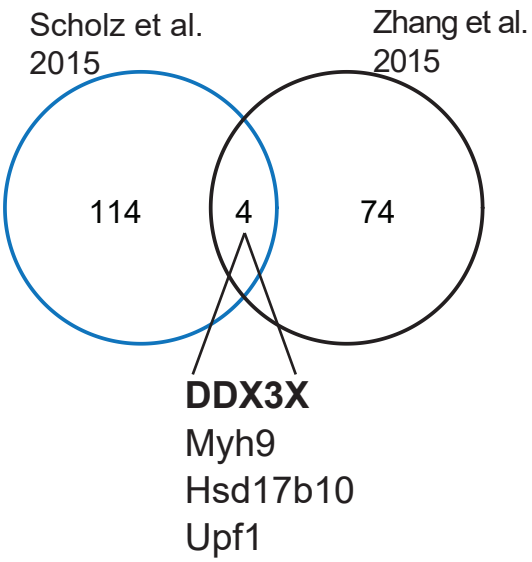


Synrg

Supplementary Figure 2. High-confidence HDAC6 target sites in IDRs (related to Figure 1).

Sequence of proteins harbouring high-confidence HDAC6 target sites were subjected to analysis for intrinsically disordered regions, using the program VSL2. The amino acid position in the linear sequence is plotted on the X-axis and the disorder tendency is displayed on the Y-axis. The position of each high-confidence HDAC6 target site in IDR is indicated and labelled in red, while it is labelled in black if it is in a structured region.

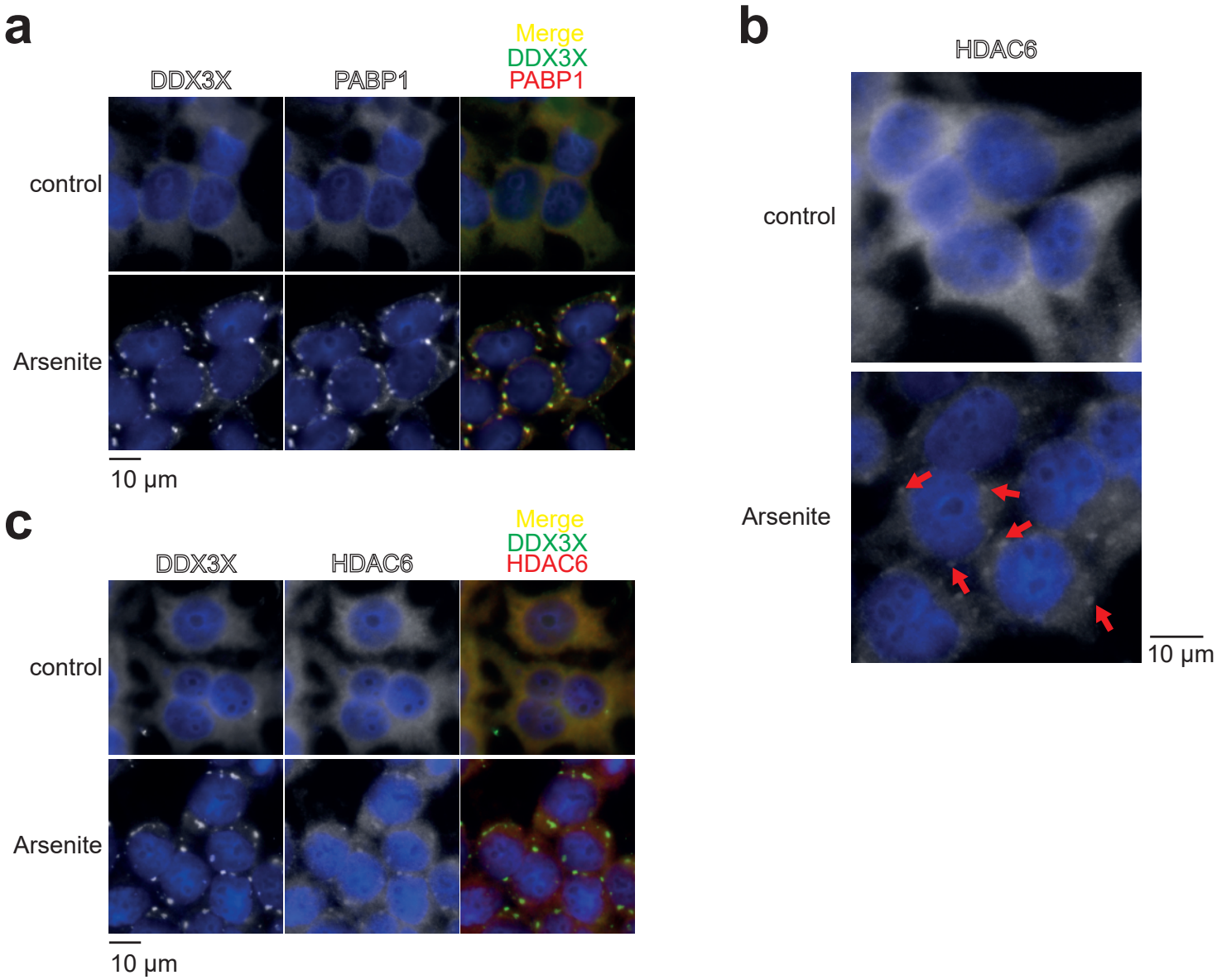
Supplementary Fig. 3



Supplementary Figure 3. Comparison of the HDAC6-dependent acetylome in two studies (related to Figure 1).

Venn diagram comparing the HDAC6-dependent acetylome in our study (based on the original acetylome of [20]) and the mouse liver HDAC6-dependent acetylome described in [24]. The number of proteins is indicated.

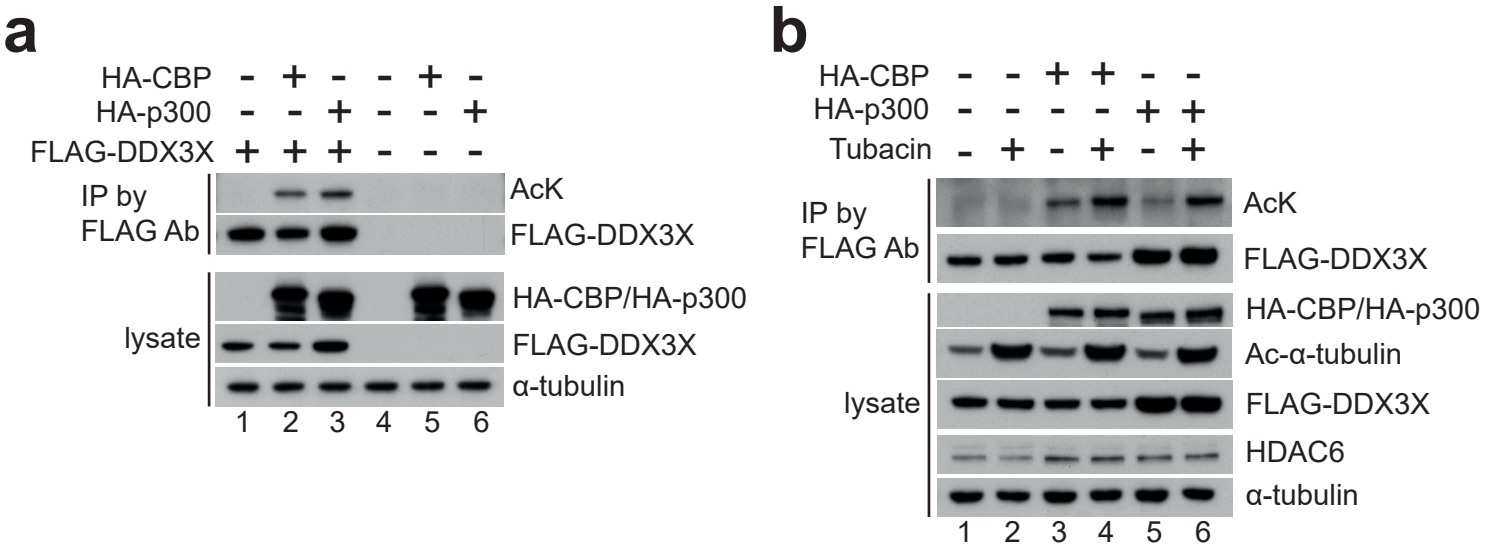
Supplementary Fig. 4

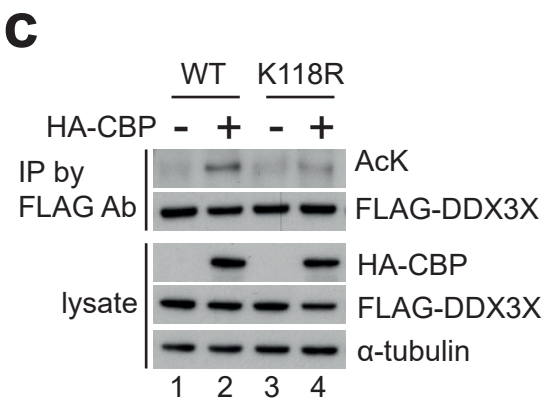


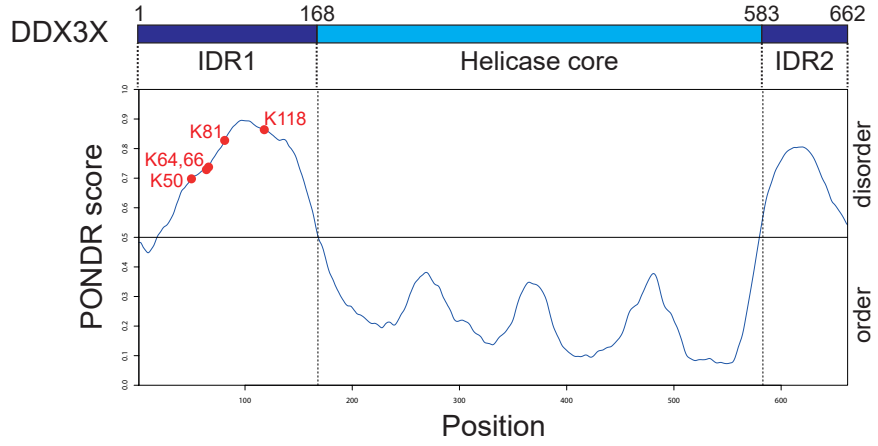
Supplementary Figure 4. Co-localization between DDX3X and HDAC6 (related to Figure 1).

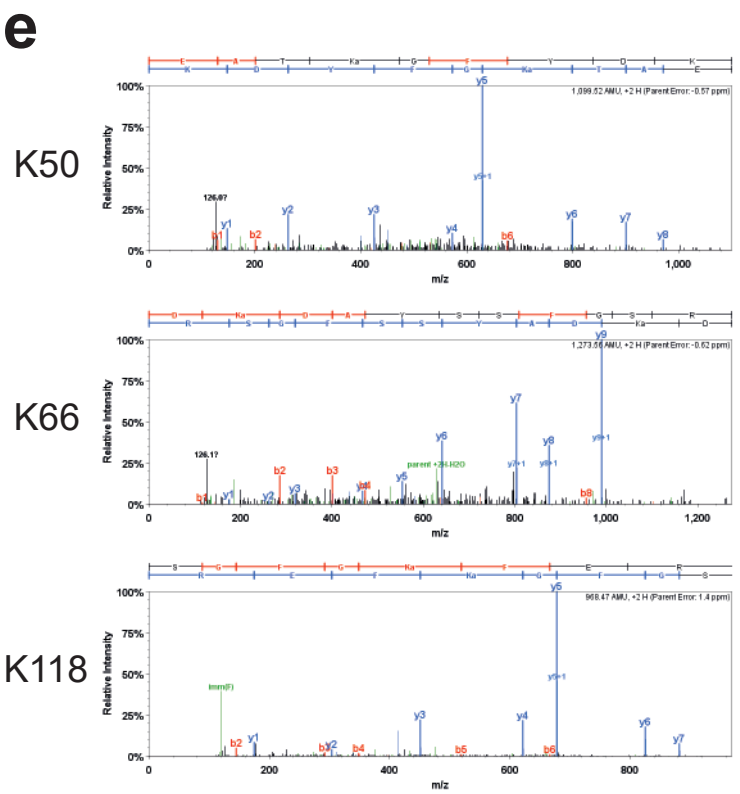
(**a**) Co-localization of DDX3X with PABP1 (stress granule (SG) marker) under oxidative stress. 293T cells were treated with 1 mM arsenite for 30 min, and localization of endogenous DDX3X and PABP1 was assessed by immunofluorescence microscopy. Representative results are shown (n = 3 biologically independent experiments). Scale bar, 10 μm. (**b**) Localization of HDAC6 on SGs under oxidative stress. 293T cells were treated with 1 mM arsenite for 30 min, and localization of endogenous HDAC6 was assessed by immunofluorescence microscopy. Red arrows indicate granules containing HDAC6. Representative results are shown (n = 3 biologically independent experiments). Scale bar, 10 μm. (**c**) Localization of DDX3X and HDAC6 on SGs under oxidative stress. 293T cells were treated with 1 mM arsenite for 30 min, and localization of endogenous DDX3X and HDAC6 was assessed by immunofluorescence microscopy. Representative results are shown (n = 3 biologically independent experiments). Scale bar, 10 μm.

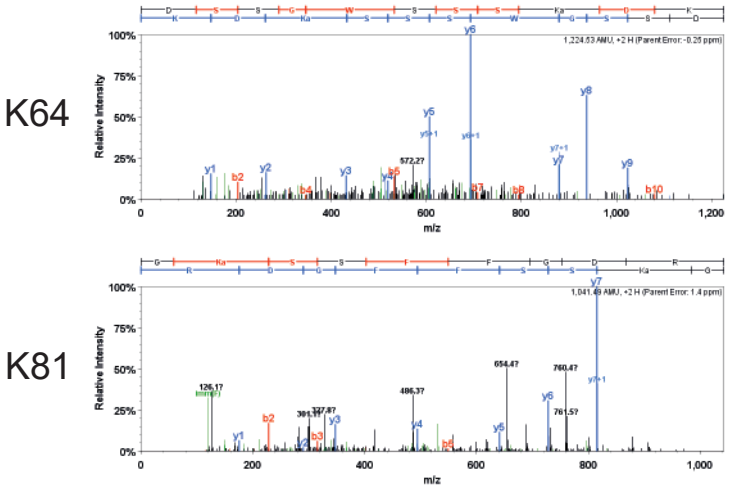
Supplementary Fig. 5









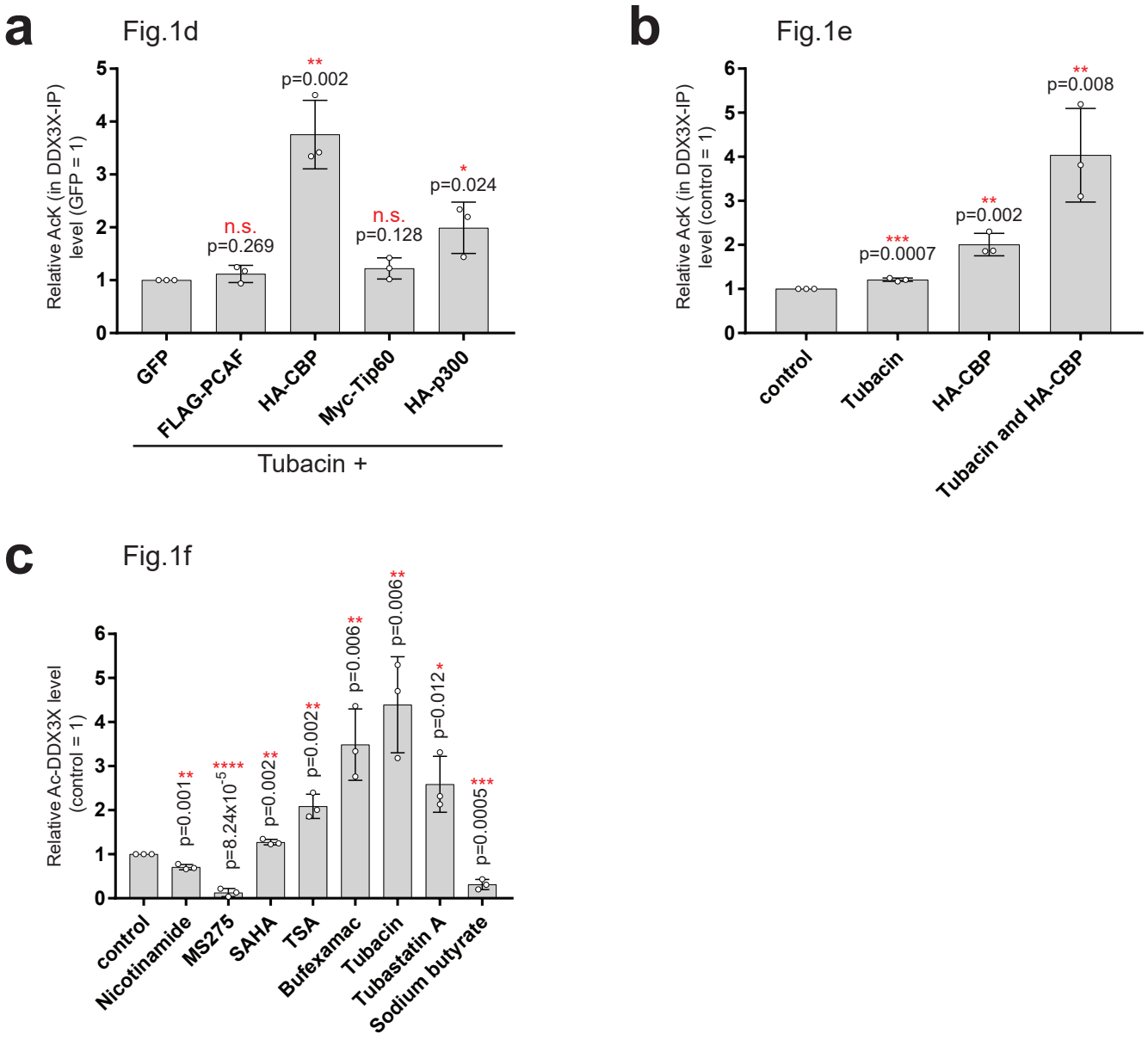


Peptide sequence	Position	Best Ascore	Best Peptide Score
EATK(Ac)GFYDK	50	48.83	78.26
DSSGWSSSK(Ac)DK	64	24.75	67.12
DK(Ac)DAYSSFGSR	66	1000	70.07
GK(Ac)SSFFGDR	81	1000	42.32
SGFGK(Ac)FER	118	1000	88.58

Supplementary Figure 5. Multiple acetylation sites are present in DDX3X-IDR1 (related to Figure 1).

(**a**) Acetylation of overexpressed DDX3X by CBP and p300. Tagged versions of DDX3X, CBP and p300 were transiently expressed in 293T cells. The acetylation status of FLAG-DDX3X was analyzed with a pan-acetyl-lysine antibody after immunoprecipitation. Representative results are shown (n = 3 biologically independent samples). Uncropped gel images can be found in **Supplementary Fig. 27i**. (**b**) Acetylation of exogenous DDX3X is increased by Tubacin treatment. 293T cells were transiently transfected with FLAG-DDX3X together with CBP or p300 and were treated overnight with 10 μM Tubacin, as indicated. Cell lysates were used to examine the acetylation status of FLAG-DDX3X by immunoblotting with a pan-acetyl-lysine antibody after immunoprecipitation. Representative results are shown (n = 3 biologically independent samples). Uncropped gel images can be found in **Supplementary Fig. 27j**. (**c**) Confirmation of the DDX3X acetylation site by immunoblotting. FLAG-DDX3X WT or K118R was expressed in 293T cells together with CBP. The acetylation status of FLAG-DDX3X was analyzed with a pan-acetyl-lysine antibody after immunoprecipitation. Representative results are shown (n = 3 biologically independent samples). Uncropped gel images can be found in **Supplementary Fig. 27k**. (**d**) Structural prediction analysis of DDX3X showing the positions of acetylated lysines. FLAG-DDX3X was immunoprecipitated with a FLAG (M2) antibody from lysates of 293T cells transfected with FLAG-DDX3X and CBP. The protein pellets were cleaved with Lys-C and trypsin, and the extracted peptides were analyzed by mass spectrometry. All identified acetylation sites were in IDR1. (**e**) MS/MS spectrum of the DDX3X acetylated peptides identified in **d**. The results were validated with the program Scaffold (version: 4.4) and Scaffold PTM (version: 2.2). Representative results are shown (n = 3 biologically independent samples).

Supplementary Fig. 6

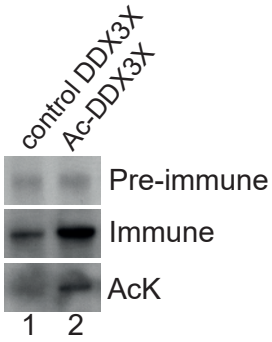


Supplementary Figure 6. Quantification of immunoblots to show DDX3X as a novel HDAC6 substrate (related to Figure 1).

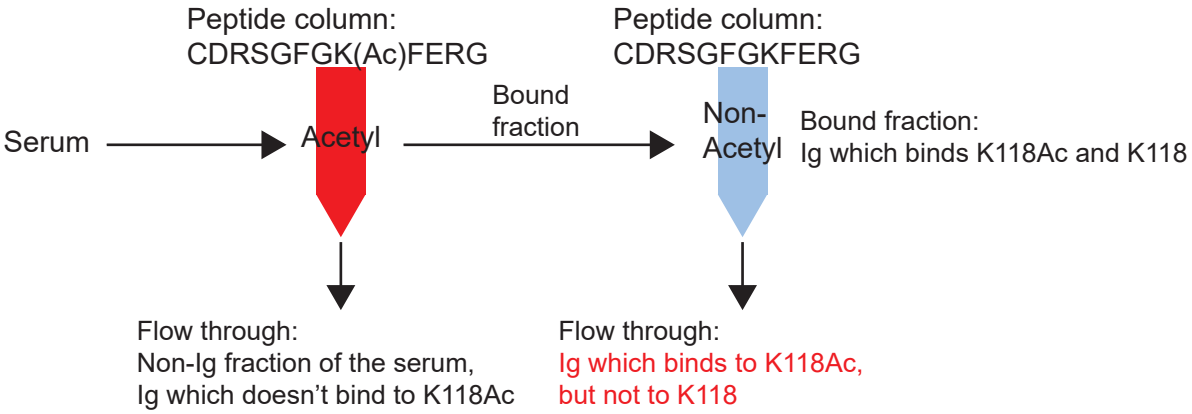
Quantification of Ac-DDX3X signals as in **Fig.1**. Intensity of the band under each condition was quantified using ImageJ, and normalized by that of control condition. Mean and SD values are presented (n = 3 biologically independent samples). P values were determined by Student's two-tailed t-test; (*) P < 0.05, (**) P < 0.01, (***) P < 0.001 and (****) P < 0.0001. (**a**) Quantification of AcK signal in DDX3X-IP product in **Fig.1d**. (**b**) Quantification of AcK signal in DDX3X-IP product in **Fig.1e**. (**c**) Quantification of Ac-DDX3X (K118) signal in **Fig.1f**.

Supplementary Fig. 7

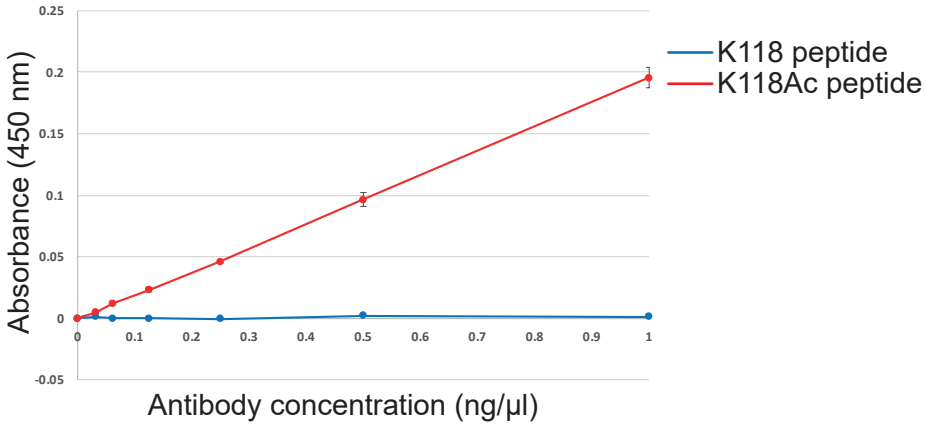
a



b



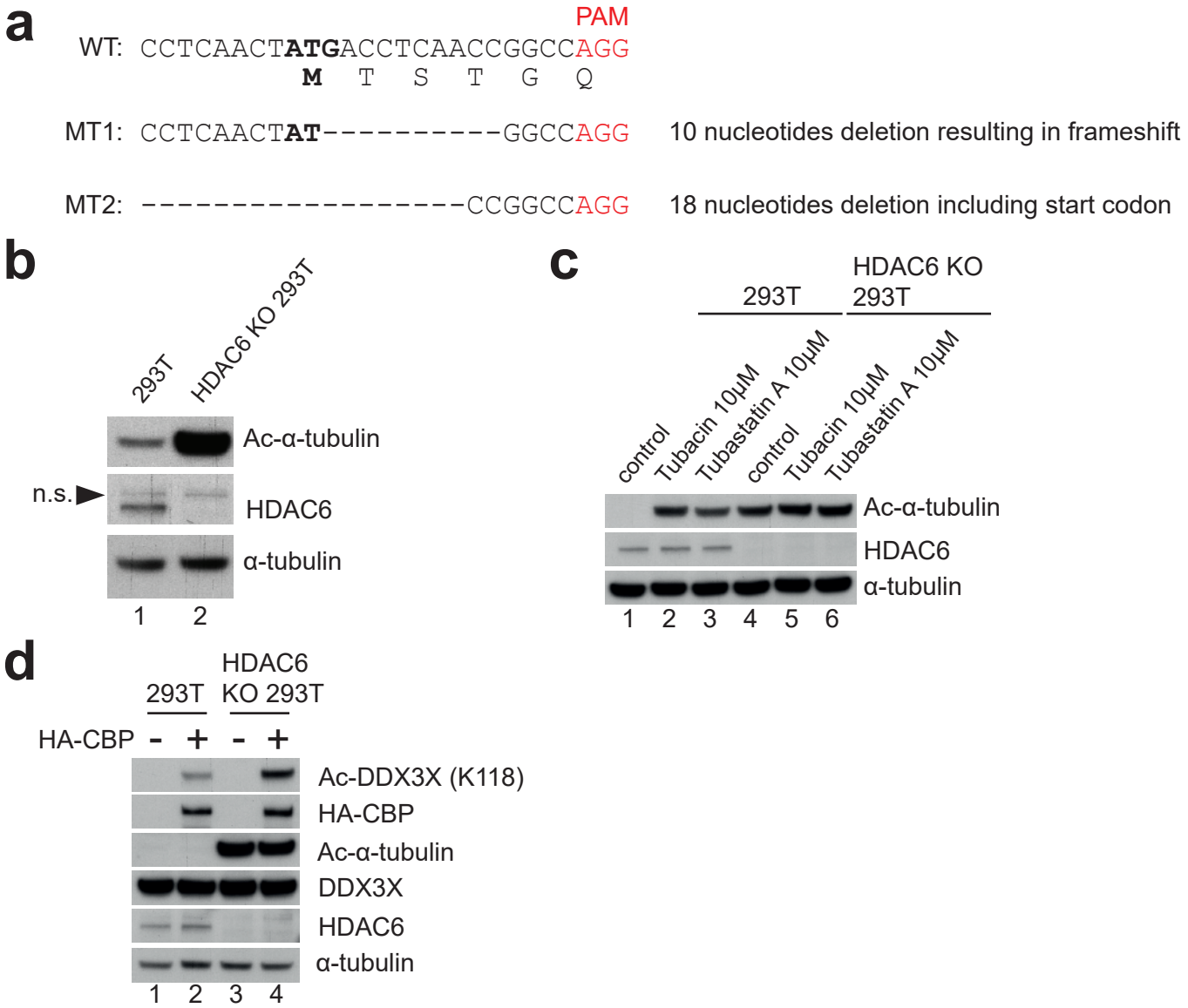
c



Supplementary Figure 7. Establishment of a DDX3X-K118Ac-specific antibody (related to Figure 1).

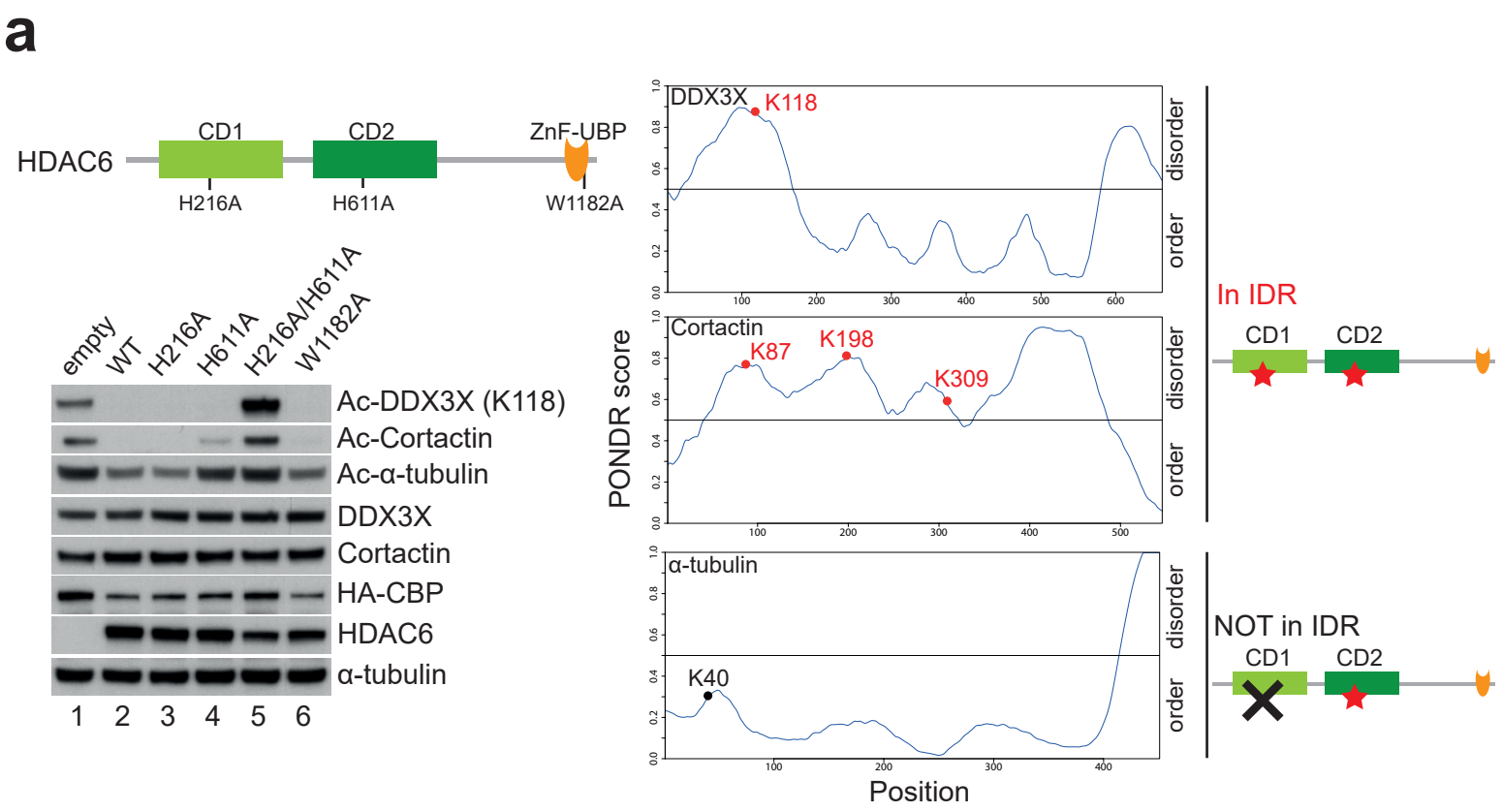
(**a**) Rabbit serum was collected 18 weeks after injection with the DDX3X-K118Ac peptide (see Methods). As a control for the analysis, 293T cells were transiently transfected with FLAG-DDX3X together with GFP (control DDX3X, lane 1) or CBP (Ac-DDX3X, lane 2) and DDX3X was immunoprecipitated from cell lysates. Each protein sample was then analyzed by immunoblotting, using pre-immune serum, immune serum or a commercial pan-acetyl-lysine antibody as indicated. Representative results are shown (n = 3 biologically independent samples). An uncropped gel image can be found in **Supplementary Fig. 27l**. (**b**) Scheme of the sequential affinity purification procedure for a DDX3X-K118Ac-specific antibody, with two columns. The serum was first subjected to affinity purification with K118Ac peptide (CSGFGK(Ac)FER) column; the bound fraction was eluted and subsequently passed over an unacetylated K118 peptide (CSGFGKFER) column, in order to remove antibodies that would also bind to the non-modified peptide. The flow through of the second column fraction contains the DDX3X-K118Ac-specific antibody and has been used in all subsequent experiments. (**c**) Specificity of DDX3X-K118Ac antibody tested by enzyme-linked immunosorbent assay (ELISA). DDX3X-K118 acetylated or unacetylated peptide-coated plates were prepared for ELISA and subsequently reacted with increasing concentration of the affinity purified antibody. Absorbance (450 nm) of ELISA substrates after the reaction was measured at each antibody concentration (mean ± SE, n = 3 independent experiments).

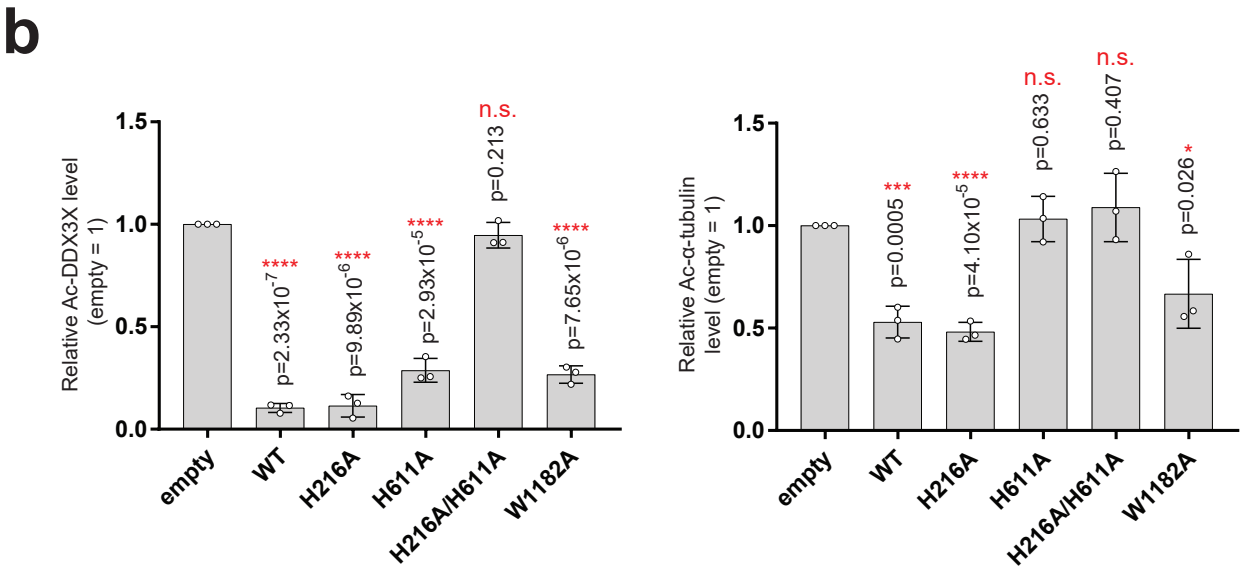
Supplementary Fig. 8



Supplementary Figure 8. Establishment of HDAC6 KO 293T (related to Figure 1).
(**a**) The top sequence (WT) shows the genomic sequence of the human HDAC6 locus around the ATG translation initiation site (black and bold) as well as the PAM site that was selected (red) for the CRISPR/Cas9 targeting. Immediately below the nucleotide sequence, the HDAC6 protein coding sequence is indicated. The two sequences below (MT1, MT2) depict the nucleotide sequence of the alleles identified in the KO cells. (**b**) Absence of HDAC6 protein and enhanced acetylation of α-tubulin in HDAC6 KO 293T cells. Expression level of HDAC6 and acetylation status of its established substrate α-tubulin were analyzed by immunoblotting in 293T and HDAC6 KO 293T cell lysates. A non-specific (n.s.) band reacting with the HDAC6 antibody is indicated by an arrowhead. Representative results are shown (n = 3 biologically independent samples). Uncropped gel images can be found in **Supplementary Fig. 27m**. (**c**) Saturated acetylation of α-tubulin in HDAC6 KO 293T cells. The acetylation status of α-tubulin was analyzed in 293T and HDAC6 KO 293T cells following treatment with HDAC6 specific inhibitors, as indicated. Representative results are shown (n = 3 biologically independent samples). Uncropped gel images can be found in **Supplementary Fig. 27n**. (**d**) Enhanced DDX3X acetylation in HDAC6 KO 293T cells. The acetylation status of endogenous DDX3X was analyzed by immunoblotting lysates from 293T and HDAC6 KO 293T cells transiently expressing CBP, as indicated; the DDX3X-K118Ac-specific antibody was used. Representative results are shown (n = 3 biologically independent samples). Uncropped gel images can be found in **Supplementary Fig. 27o**.

Supplementary Fig. 9

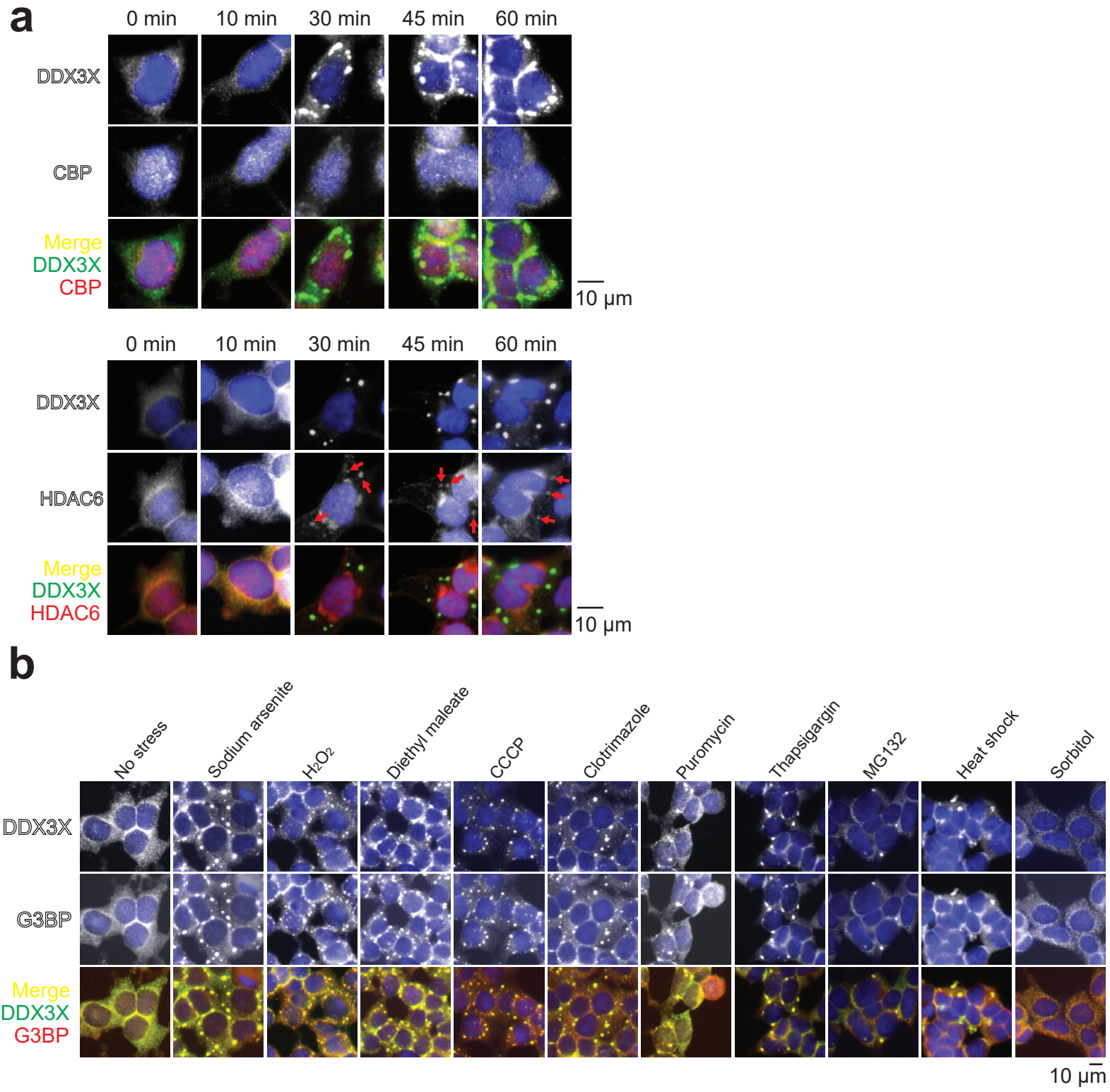




Supplementary Figure 9. DDX3X-IDR1 is a substrate of both HDAC6 CD1 and CD2 domains (related to Figure 1).

(**a**) DDX3X and cortactin can be deacetylated by either HDAC6 catalytic domain. Left, top: schematic drawing of HDAC6 structure highlighting the tandem catalytic domains (CD1, CD2) as well as the ubiquitin-binding zinc finger domain (ZnF-UBP). The positions of the different mutations are indicated. To determine which catalytic domain(s) is responsible for DDX3X-K118Ac deacetylation, different HDAC6 functional mutants (H216A, CD1 catalytic dead; H611A, CD2 catalytic dead; H216A/H611A double mutant, and W1182A, ubiquitin-binding-deficient) [27,28] were expressed in HDAC6 KO 293T cells together with CBP. Left, bottom: deacetylation of DDX3X, cortactin and α-tubulin by HDAC6 and its functional mutants. Lysates from HDAC6 KO 293T cells expressing CBP (lanes 1-6) and the different HDAC6 constructs (WT, H216A, H611A, H216A/H611A, or W1182A, lanes 2 to 6) were used to detect the acetylation status of the different proteins by immunoblotting. For DDX3X and α-tubulin, antibodies specific for the respective acetylated site were used; for cortactin a pan-acetyl-lysine antibody was used. Representative results are shown (n = 3 biologically independent samples). Uncropped gel images can be found in **Supplementary Fig. 27p**. Right: structural prediction analysis of DDX3X, cortactin and α-tubulin illustrating the position of the acetylated sites with respect to IDRs or ordered regions. The VL3-BA program was used. (**b**) Quantification of the immunoblot in **a**. Left: Ac-DDX3X (K118) signal, Right: Ac-α-tubulin signal. (mean ± SD, n = 3 biologically independent samples). P values were determined by Student's two-tailed t-test; (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$ and (****) $P < 0.0001$.

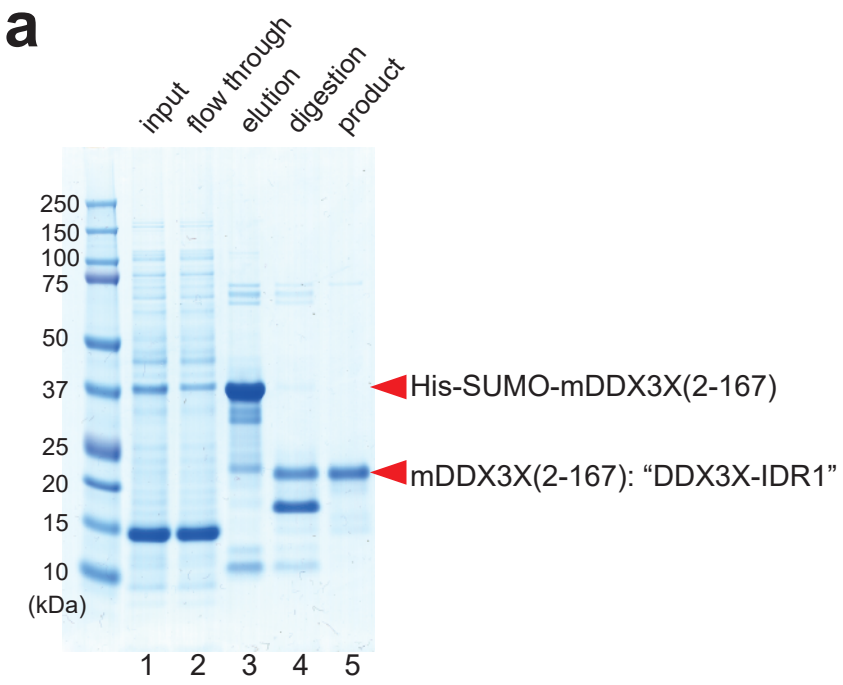
Supplementary Fig. 10

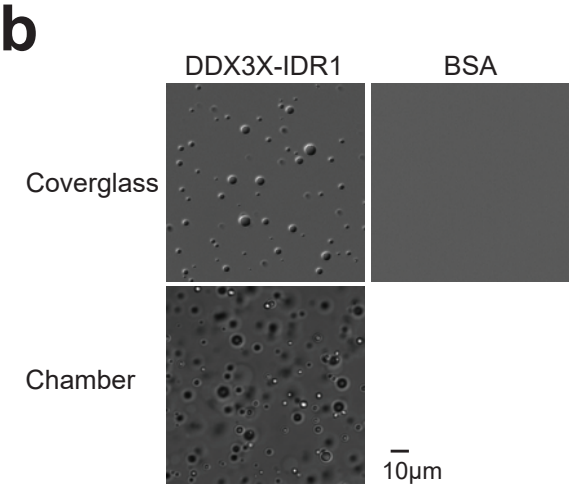


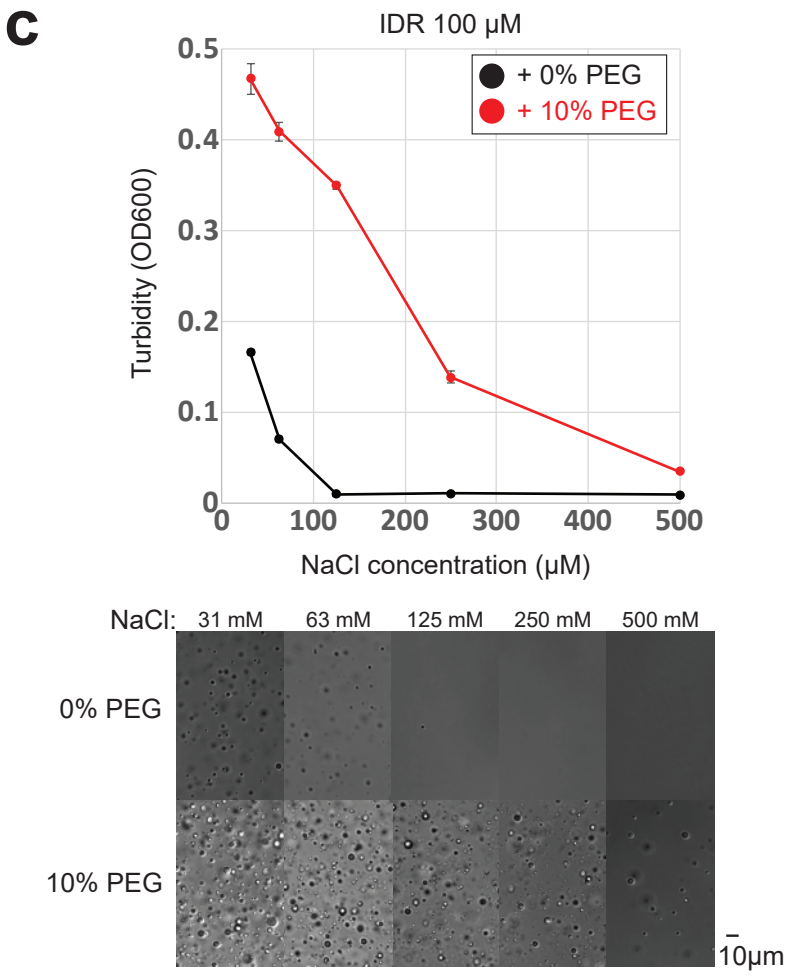
Supplementary Figure 10. Time course of SG formation in 293T under various types of stress (related to Figure 2).

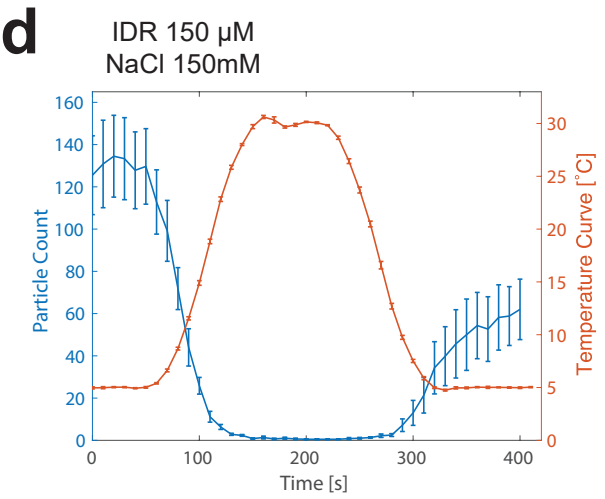
(**a**) Representative microscopy images to show the localization of DDX3X, CBP and HDAC6 in 293T at each time point after arsenite treatment. Red arrows indicate granules containing HDAC6. Representative results are shown (n = 3 biologically independent experiments). Scale bar, 10 μm. (**b**) Representative microscopy images to show the co-localization of DDX3X with G3BP (SG marker) in 293T under various types of stress. SGs were induced by the following conditions: 1 mM sodium arsenite for 1 h, 2 mM H_2O_2 for 1 h, 3 mM diethyl maleate for 1 h, 20 μM CCCP in glucose-free culture medium for 1 h, 40 μM clotrimazole in glucose-free culture medium for 3 h, 20 μg/ml puromycin for 3 h, 10 μM thapsigargin for 1 h, 10 μM MG132 for 3 h, heat shock at 43°C for 1 h and 0.5 M sorbitol for 1 h. Representative results are shown (n = 3 biologically independent experiments). Scale bar, 10 μm.

Supplementary Fig. 11



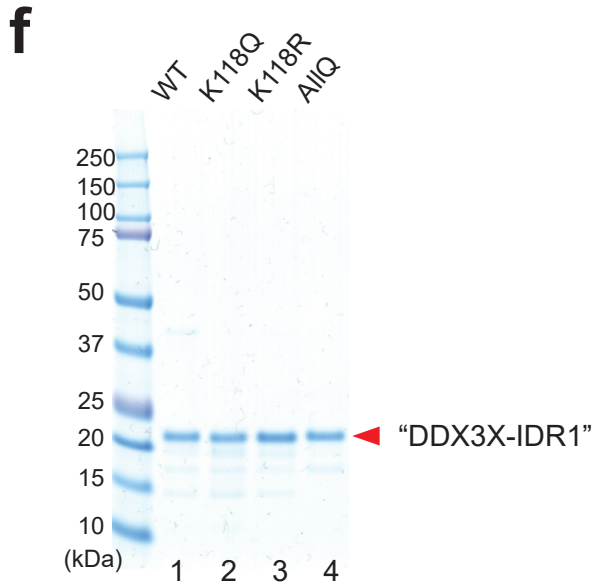






e

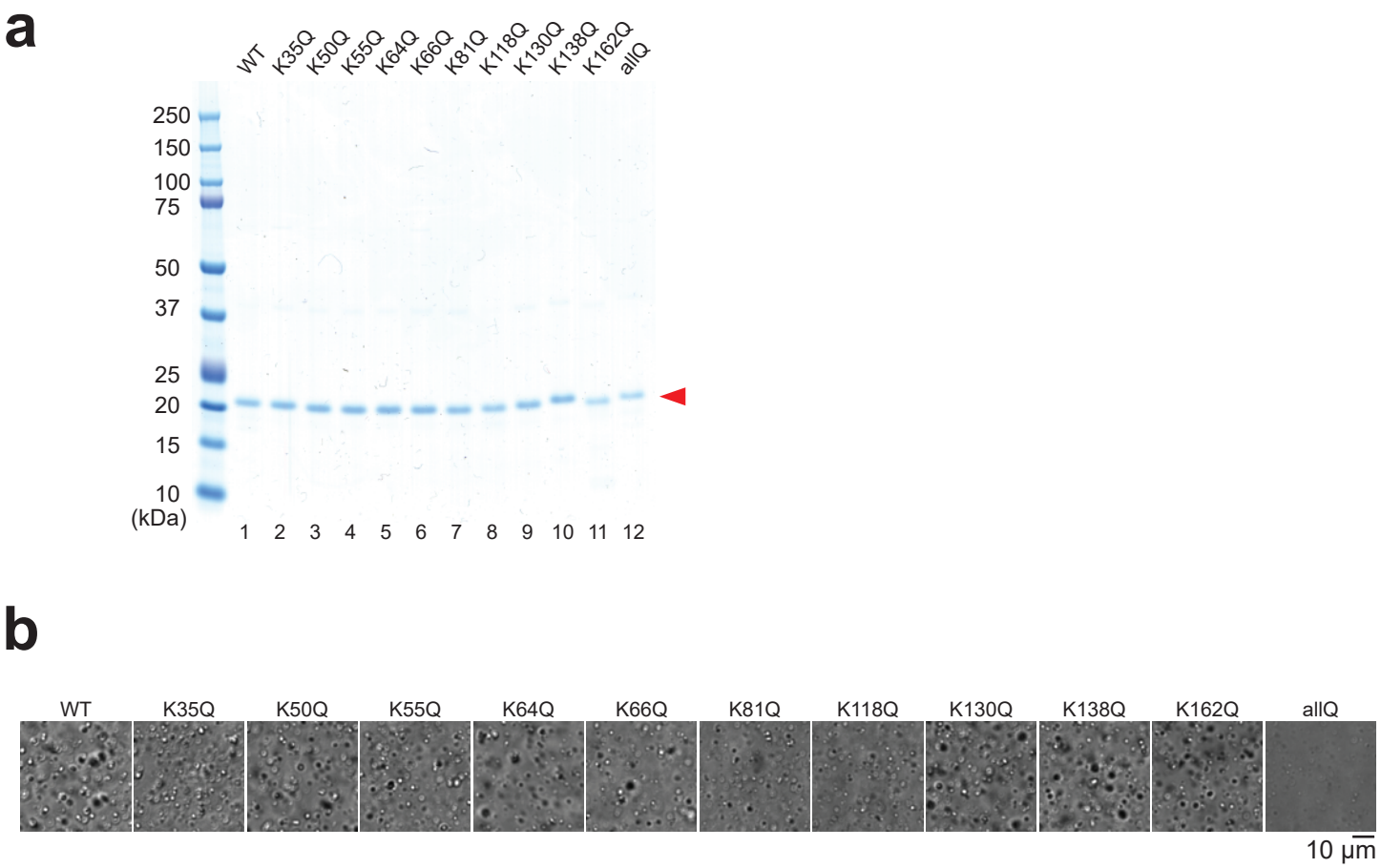
Position	Best Ascore	No reaction	Acetylation
35	1,000.00	0	108
50	1,000.00	0	69
55	1,000.00	0	17
64	1,000.00	0	26
66	1,000.00	0	5
81	1,000.00	0	143
118	1,000.00	0	80
130	1,000.00	0	104
138	1,000.00	0	11
162	1,000.00	0	18



Supplementary Figure 11. Molecular determinants of DDX3X-IDR1 phase separation (related to Figure 3).

(**a**) A Coomassie Brilliant Blue stained SDS-PAGE gel to show the process of DDX3X-IDR1 purification. The E. coli lysate was incubated with Ni-NTA resin, and bound proteins were eluted from the resin (lanes 1 to 3: input, flow through and elution). After digestion by SUMO protease, the samples were passed over Ni-NTA resin and the flow through was collected (lanes 4 and 5: digestion and product). Representative result is shown (n = 3 biologically independent samples). (**b**) Phase separation of DDX3X-IDR1. Droplet formation of 100 μM protein (BSA or DDX3X-IDR1) in 20 mM Tris, pH 7.5, 100 mM NaCl, 2 mM DTT and 10% PEG was tested. Images of formed liquid droplets at the surface of coverslip and in chamber were captured by DIC microscopy. Representative results are shown (n = 3 independent experiments). Scale bar, 10 μm. (**c**) Salt and molecular crowder effect on LLPS of DDX3X-IDR1. Turbidity (OD600) of DDX3X-IDR1 in 20 mM Tris, pH 7.5, 2 mM DTT (mean ± SD, n = 3 independent experiments) at indicated NaCl and PEG concentrations was measured (top). Images of liquid droplets in each condition were captured by DIC microscopy (bottom; Scale bar, 10 μm). (**d**) Temperature-dependent dissolution and condensation of DDX3X-IDR1 droplets. DDX3X-IDR1 (150 μM, spiked with Alexa Fluor 488 - labelled IDR1 at a molar ratio of 100:1) was dissolved in 20 mM Tris, pH 7.5, 150 mM NaCl, 2 mM DTT. The sample temperature was increased from 5°C to 30°C and decreased to 5°C, leading to dissolution and condensation of droplets. Quantification of average number of droplets formed in the images is shown (mean ± SE, n = 6 independent experiments). See also **Supplementary Movie 2**. (**e**) Multiple acetylation sites in DDX3X-IDR1 following *in vitro* reaction with CBP. DDX3X-IDR1 was incubated with CBP and acetyl-CoA for overnight at 37 °C. After protein precipitation with TCA/Acetone, samples were cleaved with Asp-N, trypsin and chymotrypsin, and the extracted peptides were analyzed by HPLC. The results were validated with the program Scaffold and Scaffold PTM. (**f**) A Coomassie Brilliant Blue stained SDS-PAGE gel showing the purified WT and mutants of DDX3X-IDR1. Representative result is shown (n = 3 biologically independent samples).

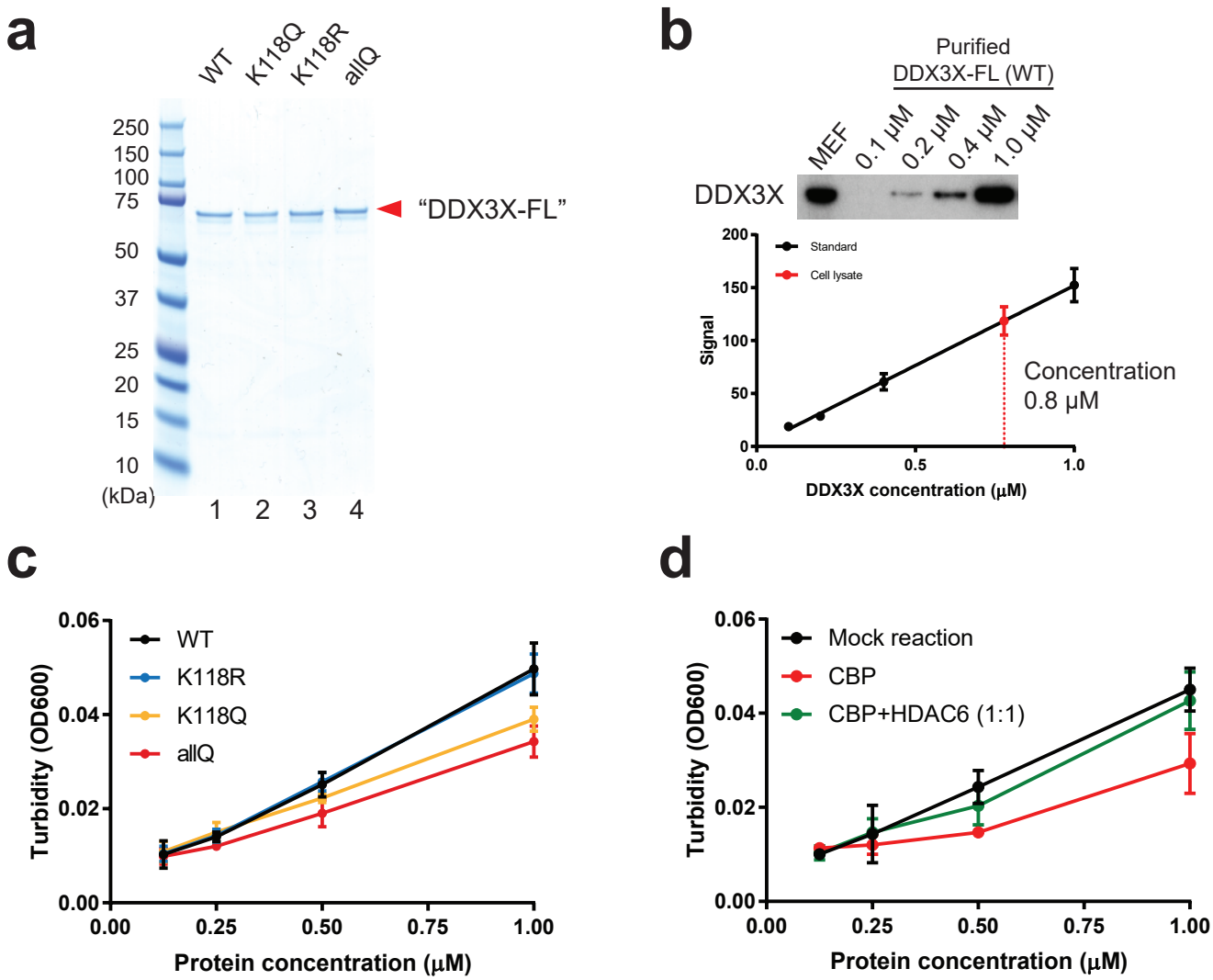
Supplementary Fig. 12



Supplementary Figure 12. Positional effects of lysine acetylation in DDX3X-IDR1 on LLPS (related to Figure 3).

(**a**) A Coomassie Brilliant Blue stained SDS-PAGE gel showing the purified WT and mutants of DDX3X-IDR1 (red arrowhead). Representative result is shown (n = 3 biologically independent samples). (**b**) Liquid droplets of WT, a series of individual K to Q, and allQ mutants of IDR1 captured by DIC microscopy. Each protein was in 20 mM Tris, pH 7.5, 200 mM NaCl, 2 mM DTT and 10% PEG. Representative results are shown (n = 3 independent experiments). Scale bar, 10 μm.

Supplementary Fig. 13



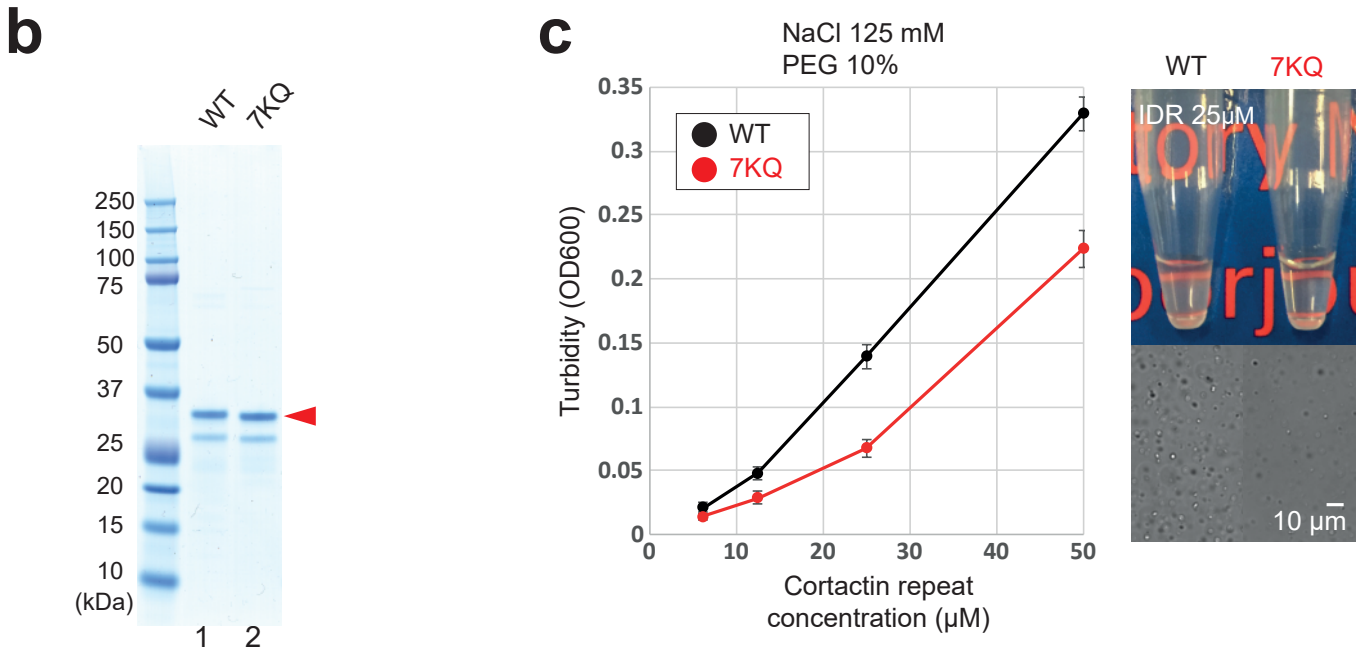
Supplementary Figure 13. Acetylation of IDR1 impairs LLPS of full-length DDX3X (related to Figure 3).

(**a**) A Coomassie Brilliant Blue stained SDS-PAGE gel showing the purified WT and mutants of DDX3X. Representative result is shown (n = 3 biologically independent samples). (**b**) Quantification of cytosolic DDX3X concentration. DDX3X in cytosolic fraction of WT MEF cells was quantified by immunoblotting, using purified full-length DDX3X as a standard. Cytosolic fraction and purified protein were equally diluted, and identical volume of samples were loaded: representative immunoblot image (top) and standard curve (bottom). DDX3X signal in immunoblot was quantified and plotted (mean ± SD, n = 3 biologically independent samples), and standard curve was obtained by linear regression. An uncropped gel image can be found in **Supplementary Fig. 27q**. (**c**) Lysine to glutamine (acetyl-lysine mimic) substitution impairs LLPS of full-length DDX3X. Turbidity (OD600) of WT, K118Q, K118R and allQ mutants of full-length DDX3X in 20 mM Tris, pH 7.5, 200 mM NaCl, 2 mM DTT and 10% PEG (mean ± SD, n = 3 independent experiments) was measured. (**d**) Recovery of LLPS by HDAC6-mediated deacetylation. Turbidity (OD600) of full-length DDX3X was measured after the reaction with CBP or CBP + HDAC6 (mean ± SD, n = 3 independent experiments).

Supplementary Fig. 14

a mCortactin repeat (a.a. 84-330)

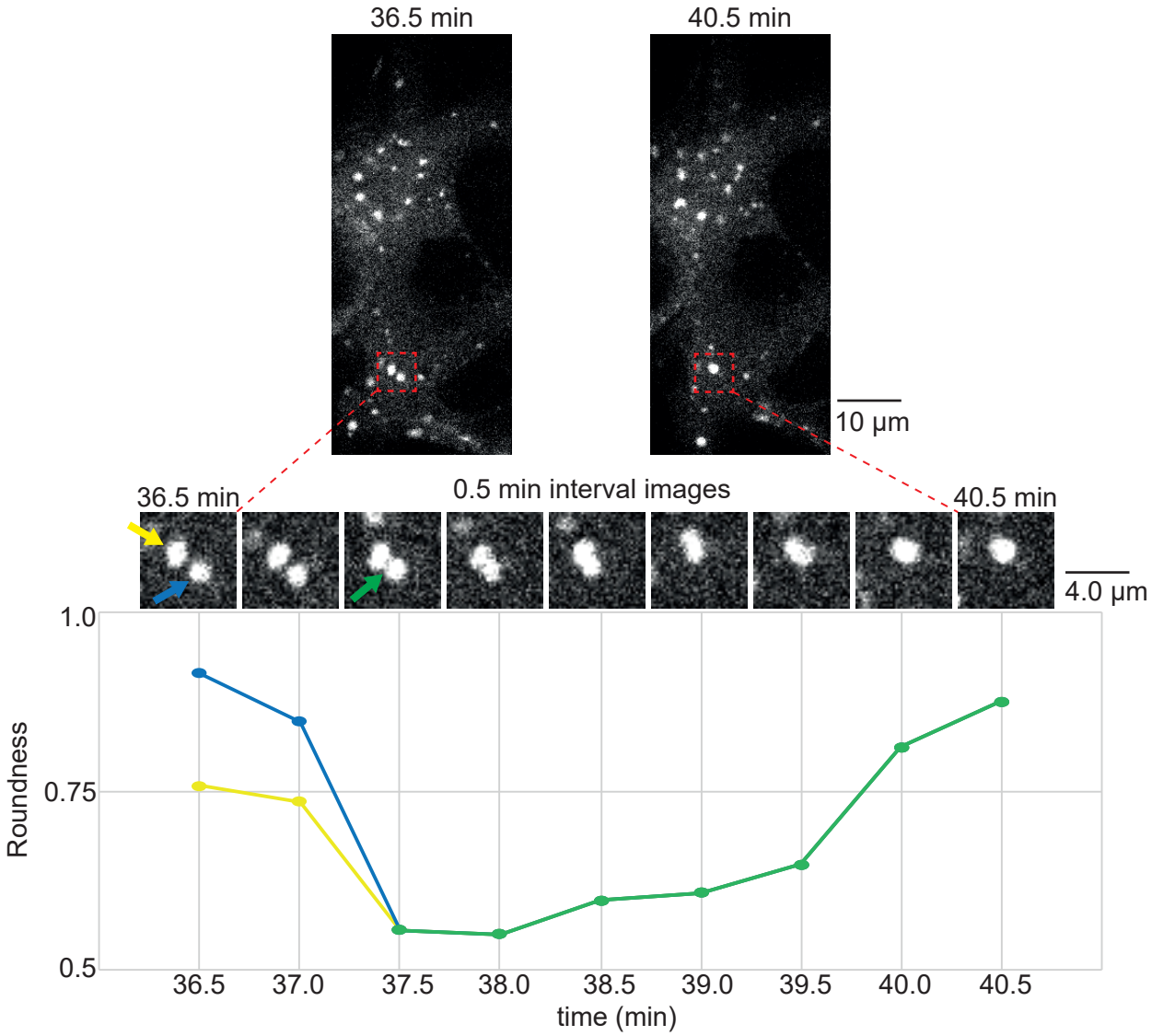
```
    YGGKFGVEQDRMDRSAVGHEYQSKLSKHCSQVD
SVRGFGGKFGVQMDRVDQSAVGFEYQGKTEKHASQKD
YSSGFGGKYGVQADRVDKSAVGFDYQGKTEKHESQKD
YSKGFGGKYGIDKDKVDKSAVGFEYQGKTEKHESQKD
YVKGFGGKFGVQTDRQDKCALGWDHQEKLQLHESQKD
YKTGFGGKFGVQSERQDSSAVGFDYKERLAKHESQQD
YAKGFGGKYGVQKDRMDKNASTFEEVVQV
```



Supplementary Figure 14. LLPS of cortactin IDR repeat is impaired by acetylation (related to Figure 3).

(**a**) Sequence of the cortactin repeat IDR (a.a. 84-330, mouse). Seven HDAC6 target GGK motifs are written in Italic and lysines within those motifs are in red. These lysines were replaced with glutamines to generate the acetyl-lysine mimic protein (7KQ). (**b**) A Coomassie Brilliant Blue stained SDS-PAGE gel showing the purified cortactin repeat (WT and 7KQ mutants, red arrowhead). Representative result is shown (n = 3 biologically independent samples). (**c**) Left: lysine to glutamine substitution impairs LLPS of cortactin repeat. Turbidity (OD600) of WT and 7KQ mutants of cortactin repeat in 20 mM Tris, pH 7.5, 125 mM NaCl, 2 mM DTT and 10% PEG (mean ± SD, n = 3 independent experiments) was measured. Right: phase separation of 25 μM cortactin repeat solutions (top) and liquid droplets captured by DIC microscopy (bottom; Scale bar, 10 μm); samples from **b**.

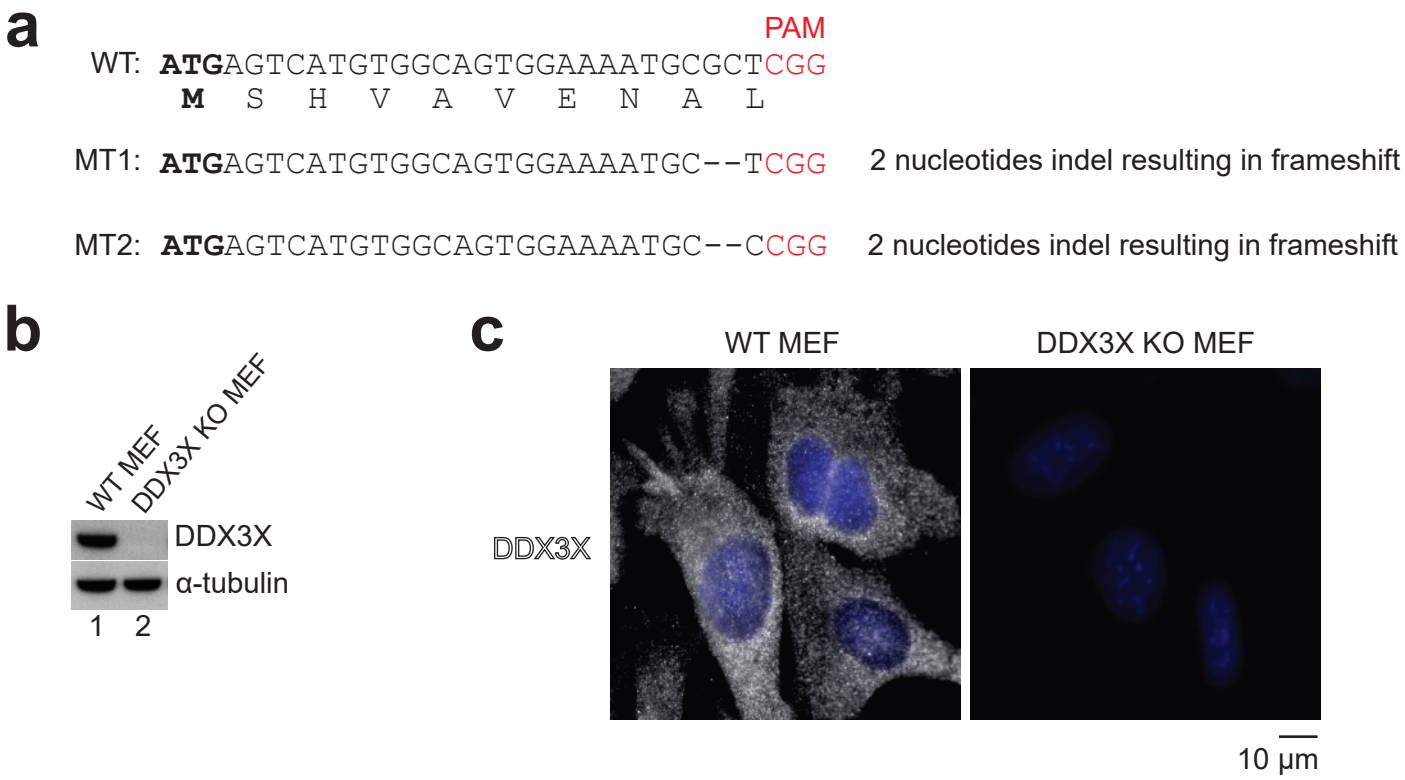
Supplementary Fig. 15



Supplementary Figure 15. Dynamic behavior of mCherry-DDX3X SGs during fusion events is consistent with LLPS (related to Figure 4).

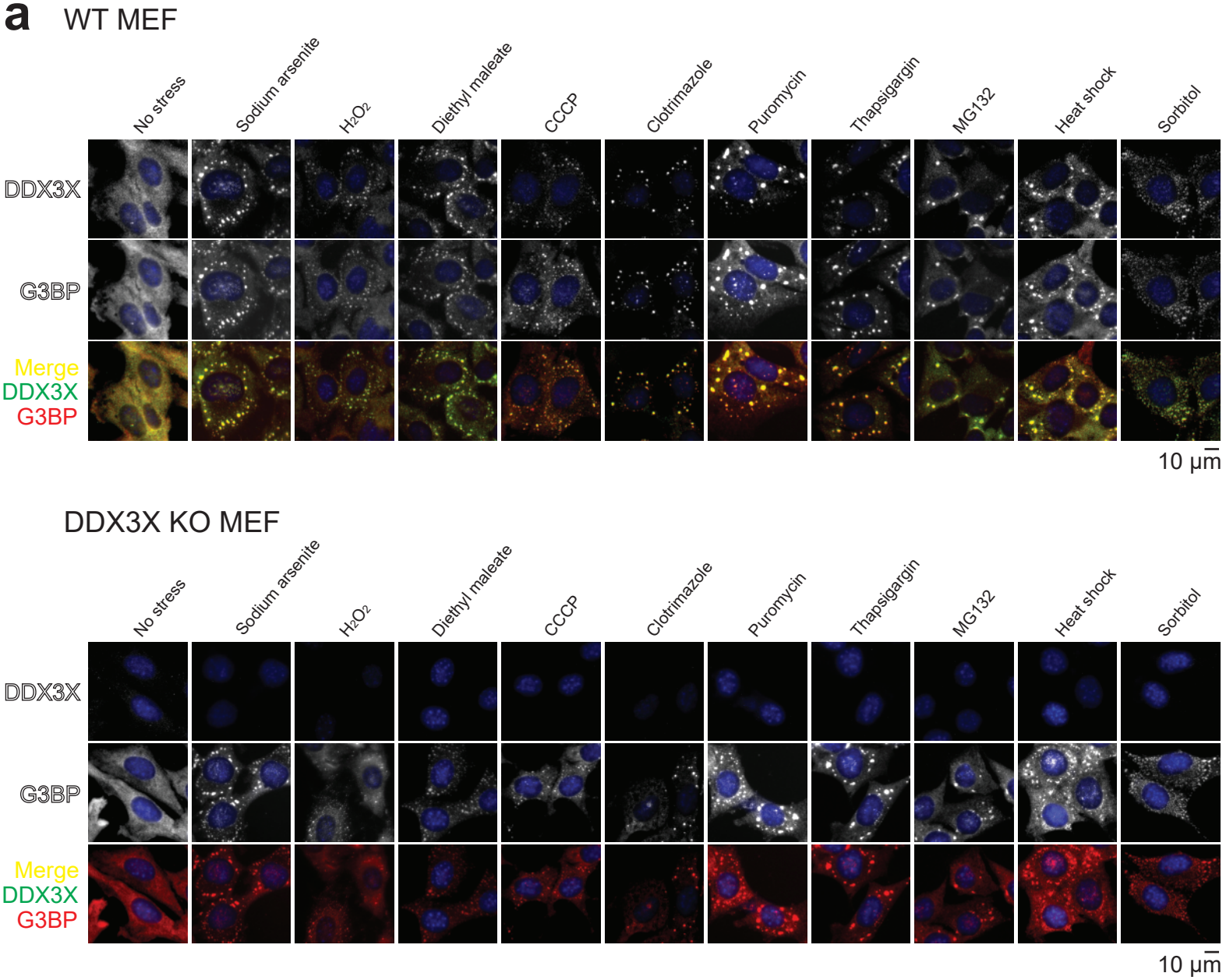
Shape of mCherry-DDX3X SGs before and after fusion process. Roundness, (4*(area)) / (π*(major axis)^2), of three granules was measured from 36.5 min to 40.5 min (time after 0.5 mM arsenite treatment, selected continuous 9 frames in **Supplementary Movie 3**). Two granules indicated by yellow and blue arrows started to fuse at 37.5 min, resulting in a new granule indicated by a green arrow. Representative result is shown (n = 3 biologically independent experiments). Scale bar, 10 μm for cell images and 4.0 μm for granule images.

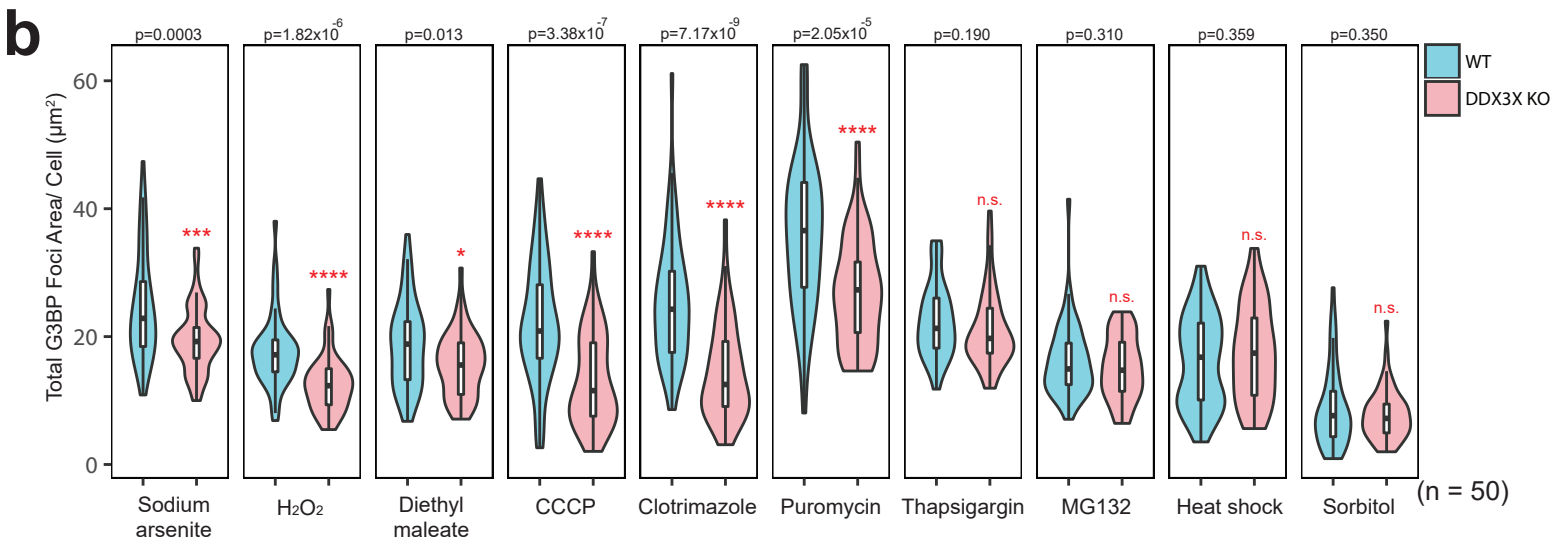
Supplementary Fig. 16



Supplementary Figure 16. Establishment of DDX3X KO MEFs (related to Figure 4). (**a**) The top sequence (WT) shows the genomic sequence of the mouse DDX3X locus around the ATG translation initiation site (black and bold) as well as the PAM site that was selected (red) for the CRISPR/Cas9 targeting. Immediately below the nucleotide sequence, the DDX3X protein coding sequence is indicated. The two sequences below (MT1, MT2) depict the nucleotide sequence of the alleles identified in the KO cells. (**b**) Absence of DDX3X protein in DDX3X KO MEFs. The expression level of DDX3X was analyzed by immunoblotting in WT and DDX3X KO MEF lysates. Representative results are shown (n = 3 biologically independent samples). Uncropped gel images can be found in **Supplementary Fig. 27r**. (**c**) Immunofluorescence microscopy analysis for the localization of endogenous DDX3X protein in WT or DDX3X KO MEFs. Representative results are shown (n = 3 biologically independent experiments). Scale bar, 10 μm.

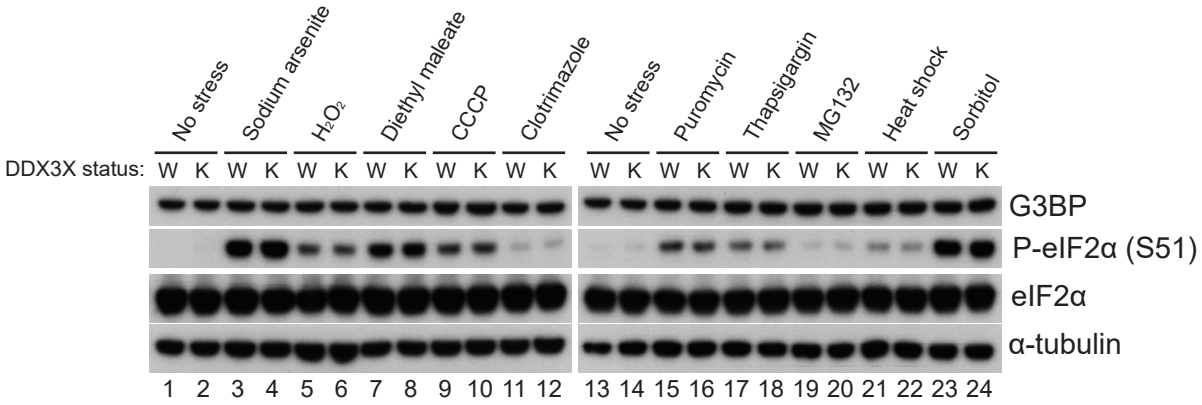
Supplementary Fig. 17_1



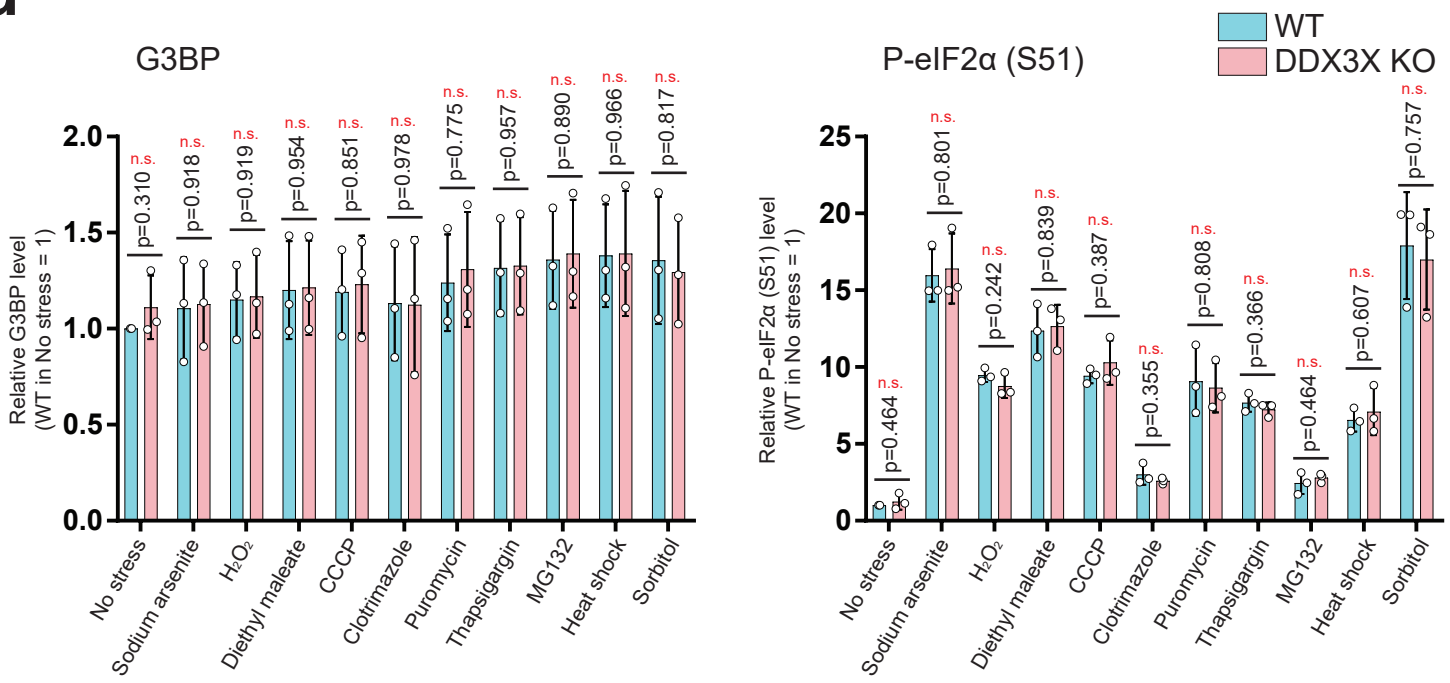


Supplementary Fig. 17_2

c



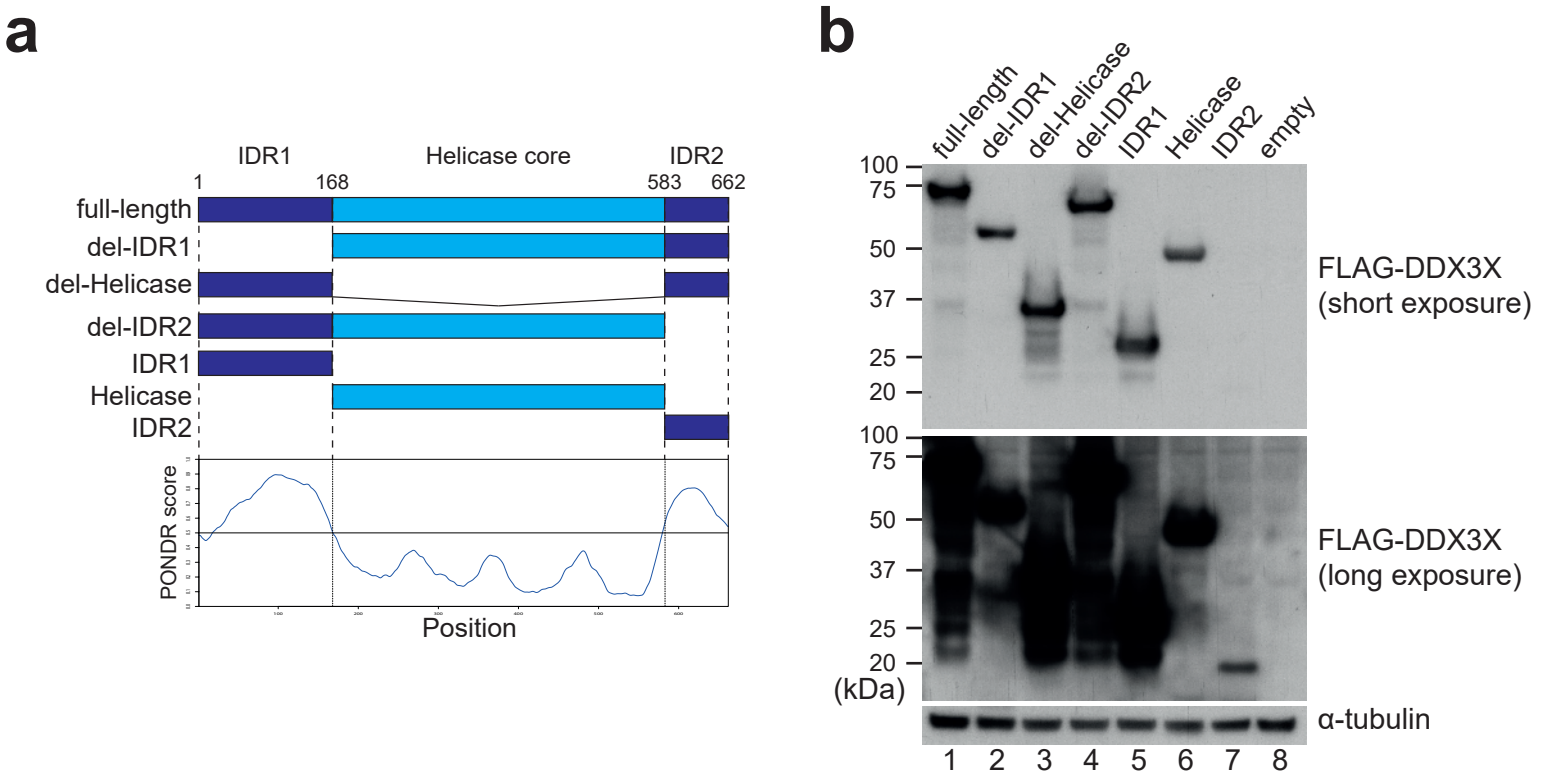
d

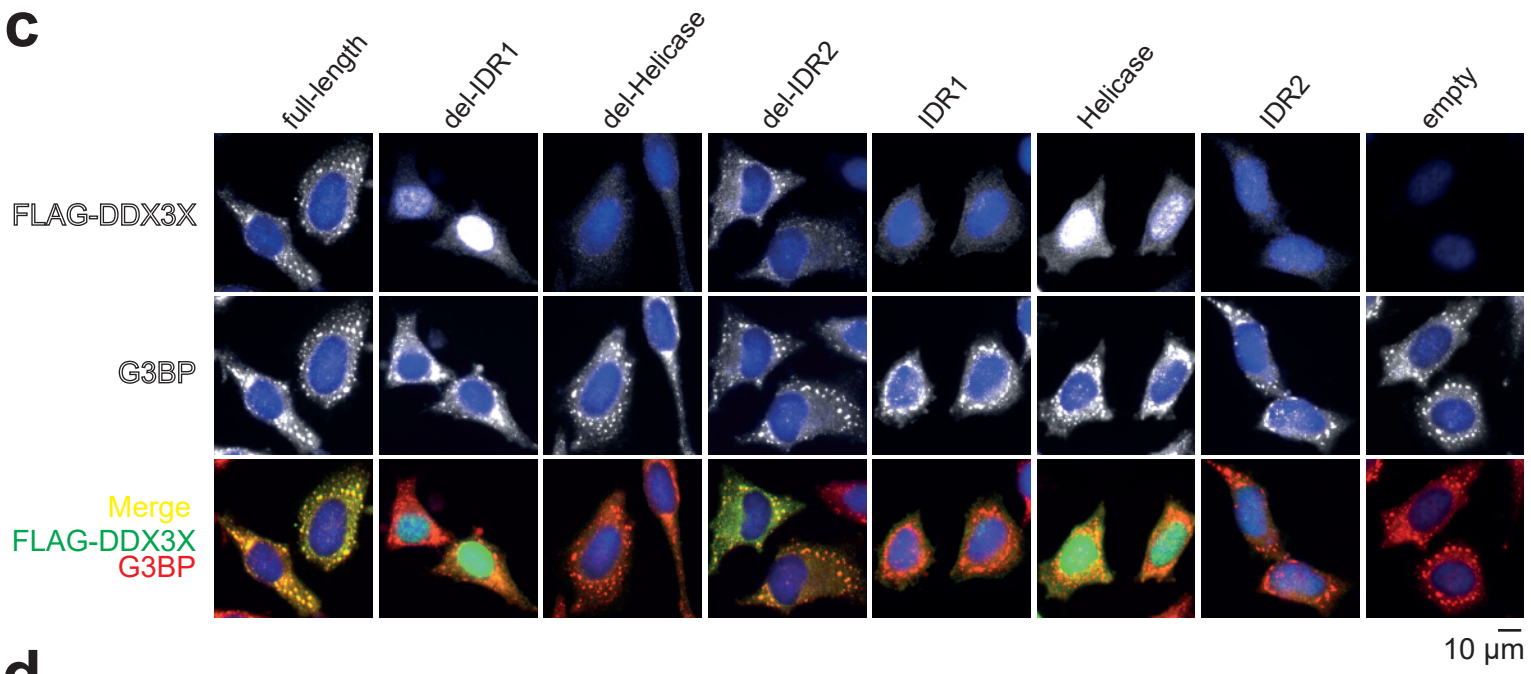


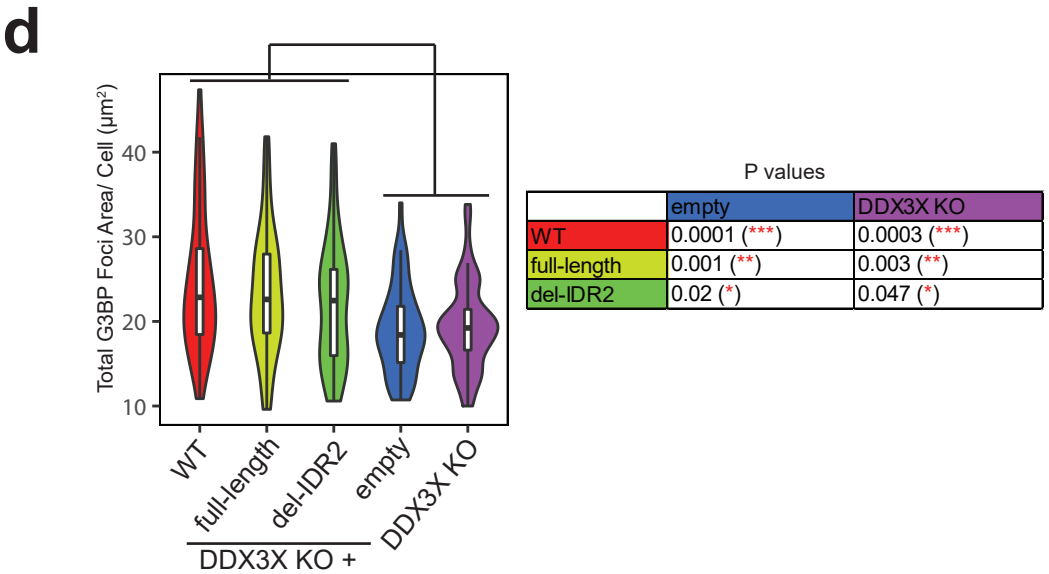
Supplementary Figure 17. DDX3X KO leads to decreased SG size under specific stress conditions (related to Figure 4).

(**a**) Representative microscopy images of SGs under various types of stresses. SGs were induced by the following conditions: 1 mM sodium arsenite for 1 h, 2 mM H_2O_2 for 1 h, 3 mM diethyl maleate for 1h, 20 μM CCCP in glucose-free culture medium for 1 h, 40 μM clotrimazole in glucose-free culture medium for 3 h, 20 μg/ml puromycin for 3 h, 10 μM thapsigargin for 1 h, 10 μM MG132 for 3 h, heat shock at 43°C for 1 h and 0.5 M sorbitol for 1 h. Localization of endogenous DDX3X and G3BP in WT and DDX3X KO MEFs was assessed by immunofluorescence microscopy. Representative results are shown (n = 3 biologically independent experiments). Scale bar, 10 μm. (**b**) Quantification of total G3BP foci area per cell (n = 50 cells in total, from biologically independent experiments) from **a**, and displayed in violin plot; the thin line at the center represents the 95% confidence interval, and the bottom and top of the box represent the 25th and 75th percentile of the data, with the median value represented by a horizontal bold line. On each side of the thin line is a kernel density estimation showing the distribution shape of the data points from minima to maxima. P values were determined by Student's two-tailed t-test; (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$ and (****) $P < 0.0001$. (**c**) G3BP expression and phosphorylation of eIF2α in DDX3X KO MEFs. Cell extracts from WT (W) and DDX3X KO (K) MEFs under each stress were analyzed by immunoblotting with each indicated antibody. Representative results are shown (n = 3 biologically independent samples). Uncropped gel images can be found in **Supplementary Fig. 27s**. (**d**) Quantification of G3BP signal and P-eIF2α signal in **c**. Mean and SD values are presented (n = 3 biologically independent samples). P values were determined by Student's two-tailed t-test; $P > 0.05$.

Supplementary Fig. 18







P values

	empty	DDX3X KO
WT	0.0001 (***)	0.0003 (***)
full-length	0.001 (**)	0.003 (**)
del-IDR2	0.02 (*)	0.047 (*)

Supplementary Figure 18. Contribution of DDX3X-IDR1 to SG formation (related to Figure 4)

(**a**) Schematic representation of the DDX3X deletion mutants analyzed. (**b**) Transient expression of DDX3X WT and its different deletion mutants in HeLa cells. The expression level of exogenous FLAG-DDX3X was analyzed by immunoblotting. Representative results are shown (n = 3 biologically independent samples). An uncropped gel image can be found in **Supplementary Fig. 27t**. (**c**) Localization of DDX3X WT and its domains deletion mutants transiently expressed in HeLa cells. Cells were treated with 1 mM arsenite for 30 min, localization of FLAG-DDX3X and endogenous G3BP was assessed by immunofluorescence microscopy. Representative results are shown (n = 3 biologically independent experiments). Scale bar, 10 μm. (**d**) Effect of DDX3X re-expression on G3BP foci area. DDX3X full-length or del-IDR2 mutant were transiently expressed in DDX3X KO MEFs. Cells were treated with 1 mM arsenite for 1 h, and total G3BP foci area per cell (n = 50 cells in total, from biologically independent experiments) was quantified and displayed in violin plot; the thin line at the center represents the 95% confidence interval, and the bottom and top of the box represent the 25th and 75th percentile of the data, with the median value represented by a horizontal bold line. On each side of the thin line is a kernel density estimation showing the distribution shape of the data points from minima to maxima. P values were determined by Student's two-tailed t-test; (*) $P < 0.05$, (**) $P < 0.01$ and (***) $P < 0.001$.

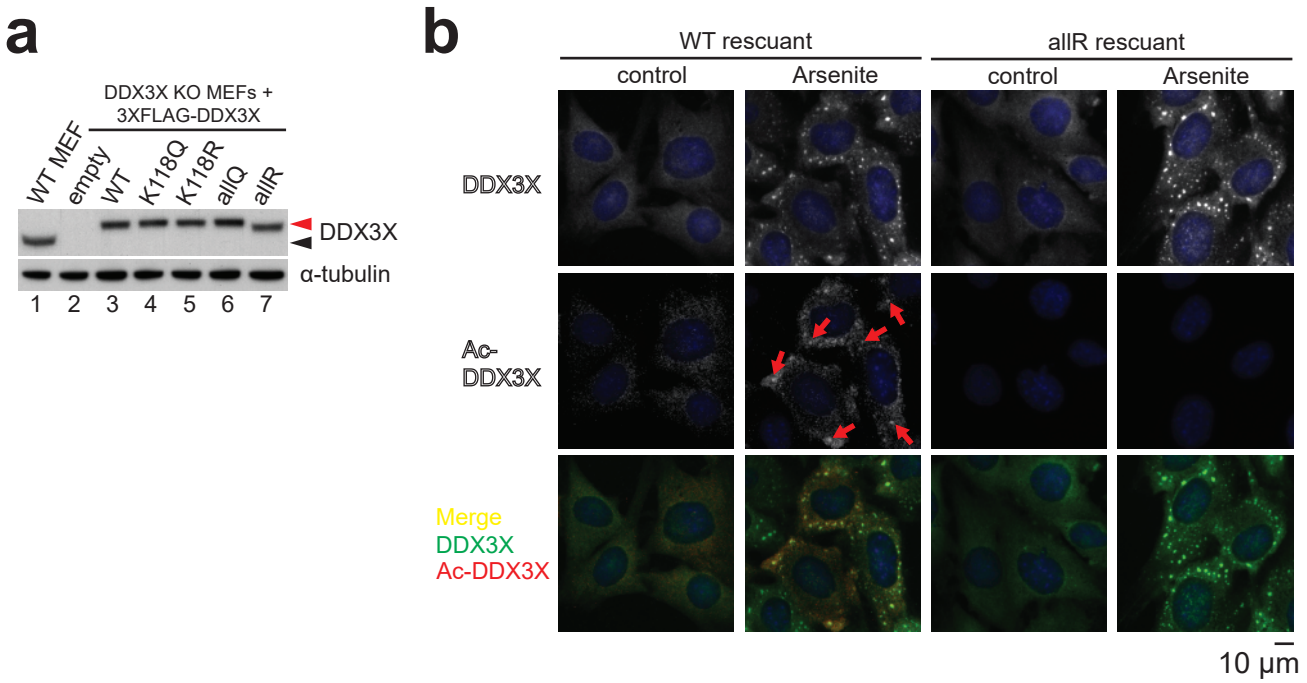
Supplementary Fig. 19

	P value (vs WT)	Halftime (s)	Mobile Fraction
allQ	3.47×10^{-36} (****)	3.087	0.8290
K118Q	0.007 (**)	5.258	0.7740
WT	-	7.324	0.7596
K118R	0.043 (*)	8.682	0.7487
allR	0.00001 (****)	10.61	0.7329

Supplementary Figure 19. Halftime and mobile fraction of mCherry-DDX3X measured by FRAP (related to Figure 4).

Quantification of fluorescence recovery after photobleaching (FRAP) of mCherry-DDX3X SGs. Halftime (seconds) and mobile fraction in **Fig. 4c**; P values were determined by one-way ANOVA; (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$ and (****) $P < 0.0001$.

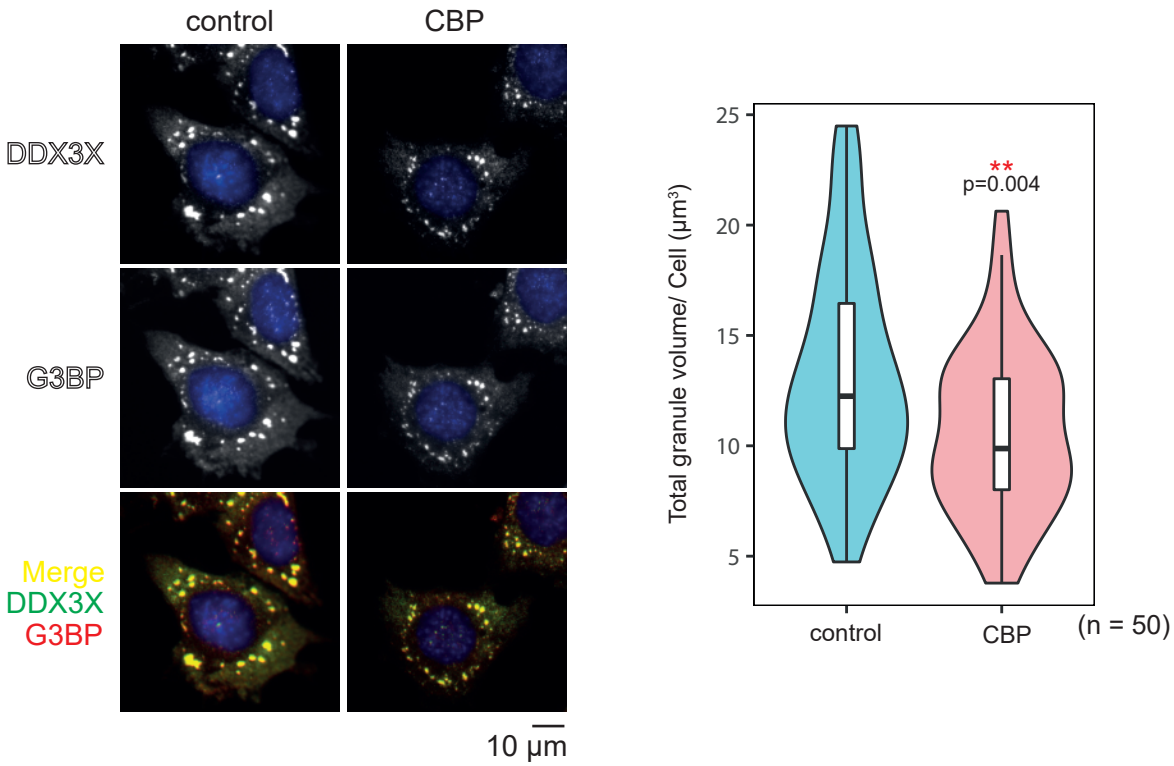
Supplementary Fig. 20



Supplementary Figure 20. Establishment of DDX3X rescuant cell lines and localization of Ac-DDX3X (related to Figure 5).

(**a**) Stable expression of DDX3X WT, K118Q/R mutants and allQ/R mutants (all lysines in IDR1 mutated to glutamine or arginine) in DDX3X KO MEFs. The expression level of endogenous DDX3X and exogenous FLAG-DDX3X was analyzed by immunoblotting with DDX3X antibody; black arrowhead indicates endogenous DDX3X while red one indicates FLAG-DDX3X. Representative results are shown (n = 3 biologically independent samples). Uncropped gel images can be found in **Supplementary Fig. 27u**. (**b**) Localization of Ac-DDX3X in DDX3X WT and allR rescuants. Cells were treated with 1 mM arsenite for 30 min, localization of FLAG-DDX3X and its K118-acetylated form were assessed by immunofluorescence microscopy. Red arrows indicate granules containing Ac-DDX3X. Representative results are shown (n = 3 biologically independent experiments). Scale bar, 10 μm.

Supplementary Fig. 21



Supplementary Figure 21. Decreased total volume of SGs upon CBP overexpression (related to Figure 5).

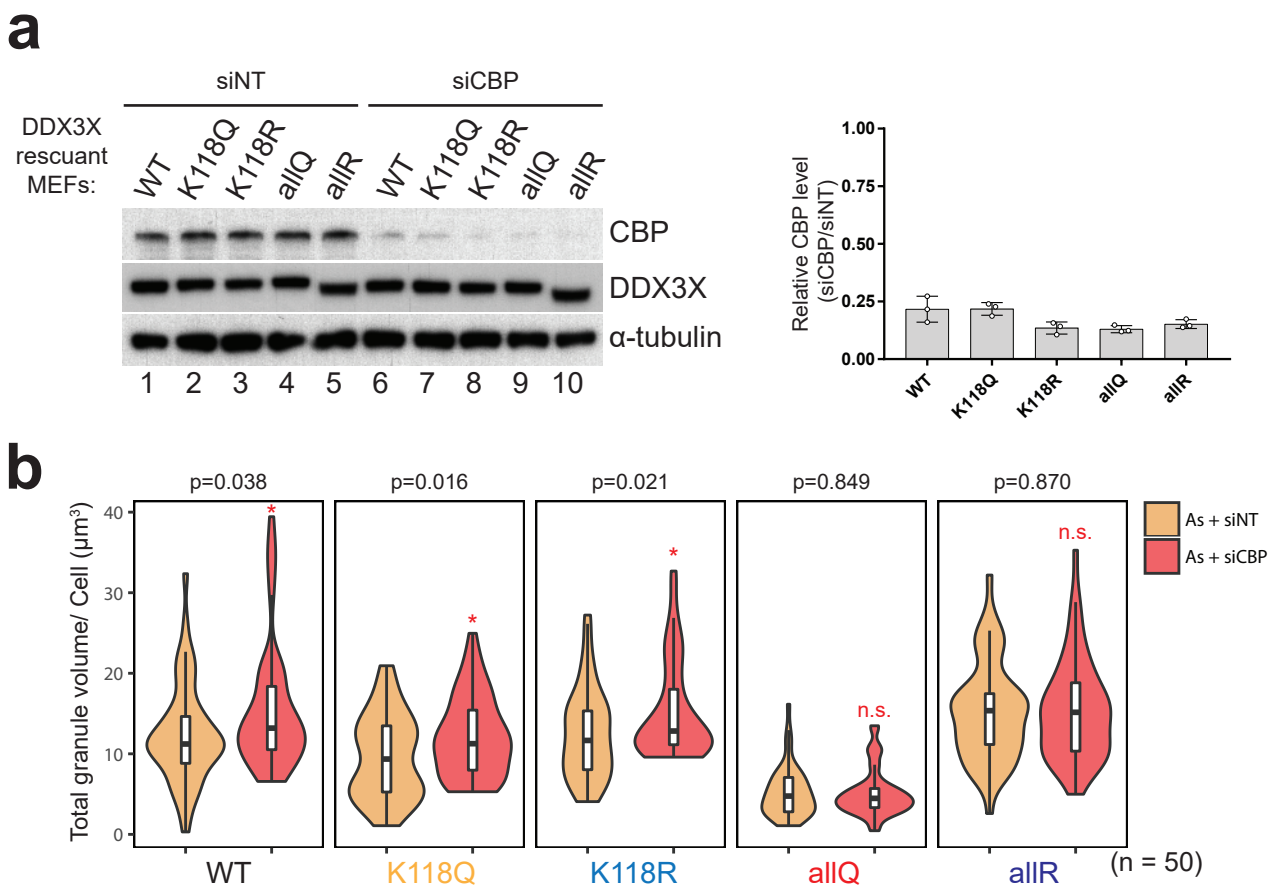
Left: immunofluorescence microscopy of SGs in WT MEF cells that had been transiently transfected with empty vector (control) or CBP in order to enhance DDX3X-IDR1 acetylation. Scale bar, 10 µm. Right panel: quantification of the number of SGs per cell (n = 50 cells in total, from biologically independent experiments) displayed in violin plot; the thin line at the center represents the 95% confidence interval, and the bottom and top of the box represent the 25th and 75th percentile of the data, with the median value represented by a horizontal bold line. On each side of the thin line is a kernel density estimation showing the distribution shape of the data points from minima to maxima. P values were determined by Student's two-tailed t-test; (**) $P < 0.01$.

Supplementary Fig. 22

Stress Type	DDX3X-dependence of SG formation	DDX3X-IDR1 acetylation by stress
Sodium arsenite	✓	✓
H_2O_2	✓	✓
Diethyl maleate	✓	✓
CCCP	✓	✗
Clotrimazole	✓	✗
Puromycin	✓	✓
Thapsigargin	✗	✗
MG132	✗	✗
Heat shock	✗	✗
Sorbitol	✗	✓

Supplementary Figure 22. Summary of DDX3X-dependence of SG formation and acetylation status of DDX3X under tested stresses (related to Figure 5).
DDX3X-dependence of SG formation is shown in **Supplementary Fig. 17**, while DDX3X-IDR1 acetylation by stress is shown in **Fig. 2c** and **d**.

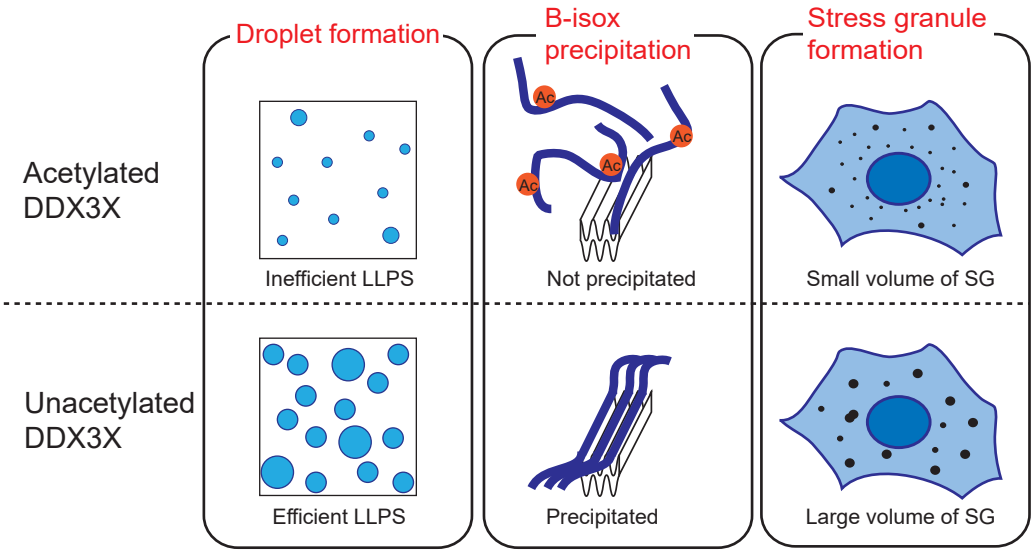
Supplementary Fig. 23



Supplementary Figure 23. Effects of CBP knockdown on DDX3X SG formation (related to Figure 5).

(**a**) Knockdown of CBP with siRNA. Left: Each DDX3X rescuant cell line (DDX3X KO MEFs stably expressing FLAG-DDX3X WT or mutants) was transfected with siRNA targeting CBP (or non-targeting "siNT" as a control) by nucleofection. CBP depletion was confirmed after 24 h of siRNA treatment as shown by immunoblotting analysis with specific antibodies. Right: Quantification of CBP signal. Each signal of siCBP-treated sample was normalized by that of siNT-treated sample (mean ± SD, n = 3 biologically independent samples). Uncropped gel images can be found in **Supplementary Fig. 27v**. (**b**) DDX3X SG volume after CBP knockdown. After 24 h of siRNA treatment, each DDX3X rescuant cell lines were treated with 1 mM arsenite for 1 h, and localization of DDX3X was assessed by immunofluorescence microscopy. Total volume of DDX3X SG per cell (n = 50 cells in total, from biologically independent experiments) was quantified and displayed in violin plot; the thin line at the center represents the 95% confidence interval, and the bottom and top of the box represent the 25th and 75th percentile of the data, with the median value represented by a horizontal bold line. On each side of the thin line is a kernel density estimation showing the distribution shape of the data points from minima to maxima. P values were determined by Student's two-tailed t-test; (*) $P < 0.05$.

Supplementary Fig. 24

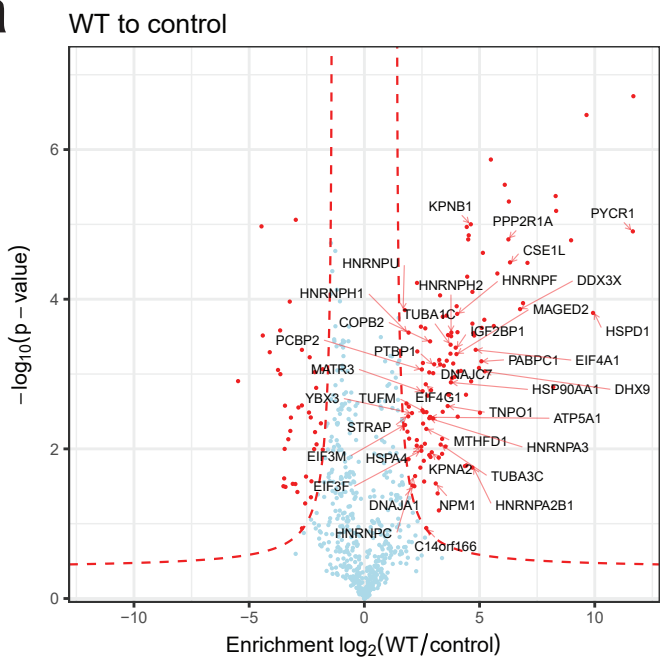


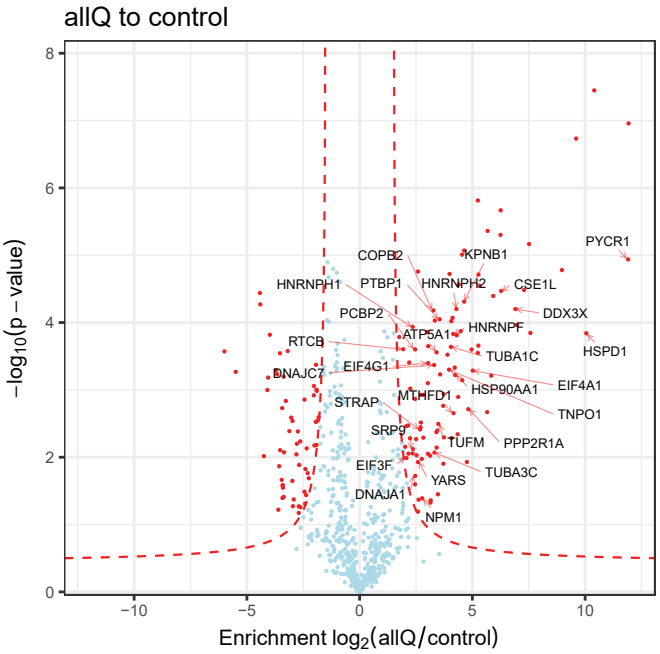
Supplementary Figure 24. Schematic representation of the observations in this study (related to Figure 5).

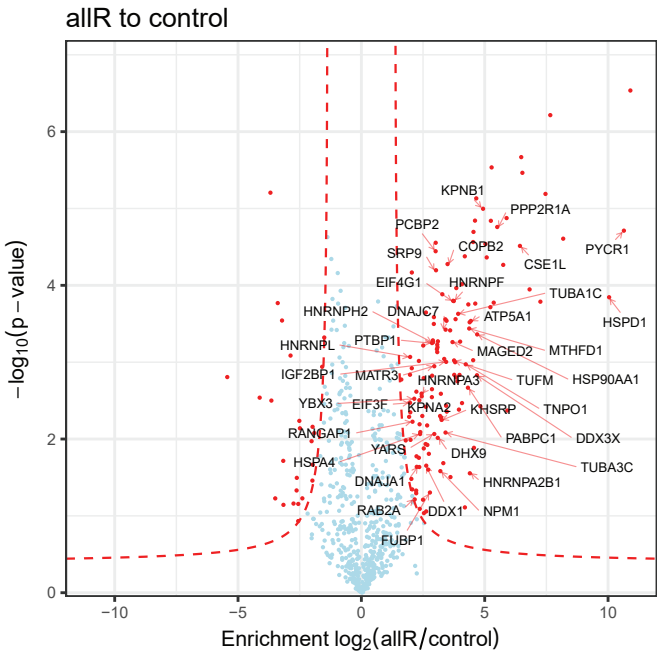
DDX3X N-terminal IDR (IDR1) forms liquid droplets through LLPS, and is efficiently precipitated *in vitro* by b-isox, which selectively promotes the aggregation of IDR-containing proteins. Acetylation of the IDR1 impairs its liquid droplet formation and changes the b-isox aggregation propensity *in vitro*. In cells, elevated acetylation of DDX3X due to HDAC6 inhibition or CBP overexpression leads to small total volume of SGs, while unacetylated DDX3X forms large volume of SGs.

Supplementary Fig. 25_1

a



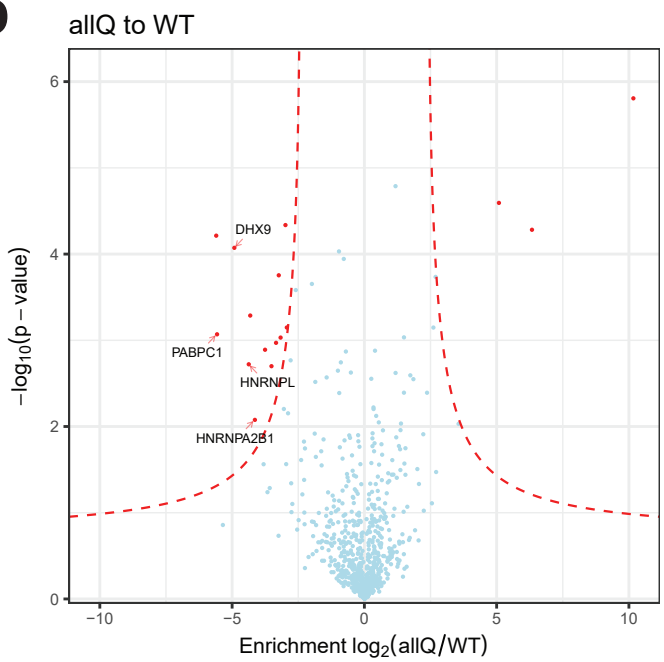


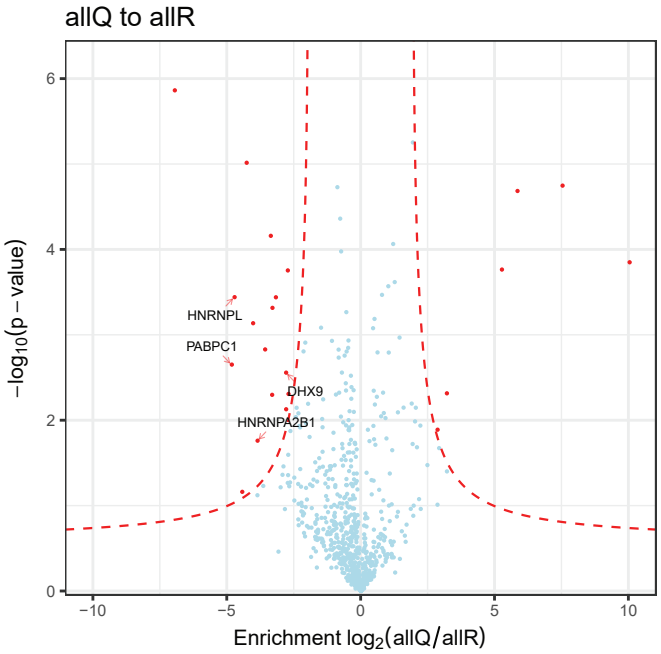


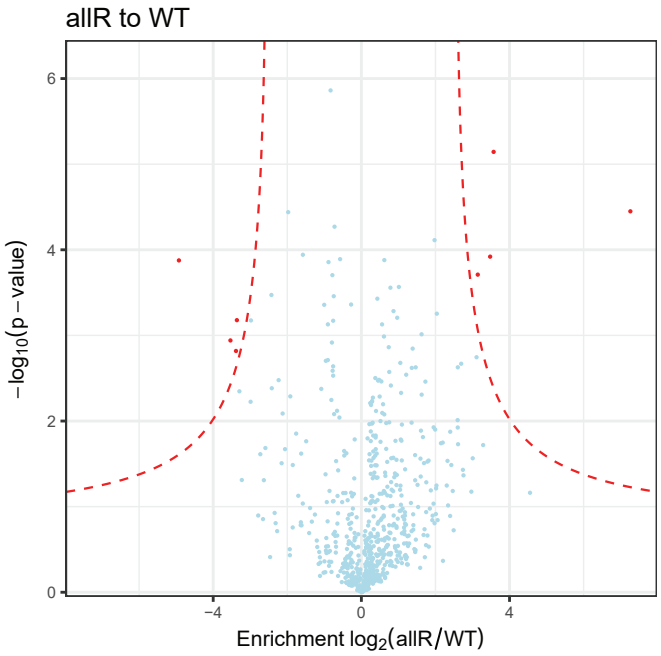
FDR = 0.03, S0 = 1.5

Supplementary Fig. 25_2

b







FDR = 0.03, S0 = 1.5

Supplementary Figure 25. Interactome analysis of DDX3X WT and acetyl-mimic/dead mutants (related to Figure 6).

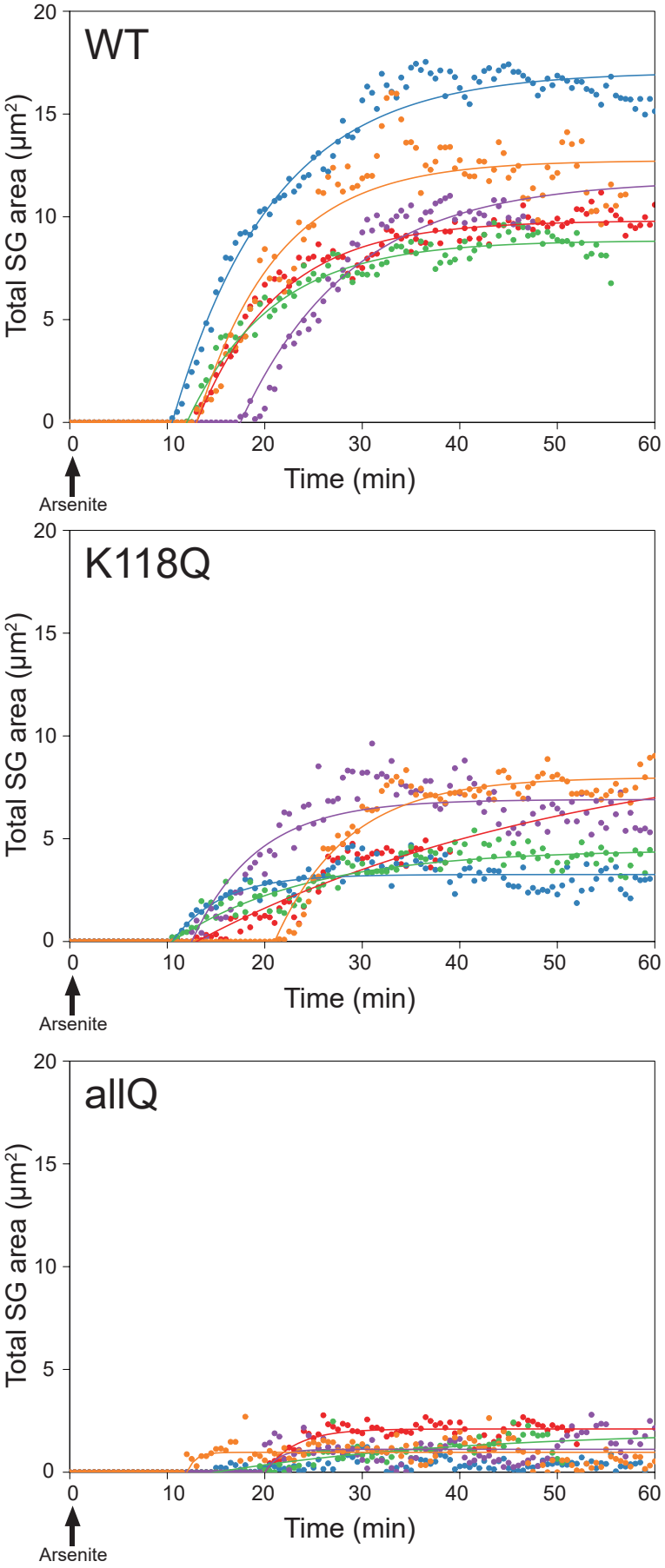
(**a**) Interactomes of each DDX3X WT and acetyl-mimic/dead mutant. Each FLAG-DDX3X construct was expressed in 293T cells, immunoprecipitated with FLAG (M2) antibody, and their interactors were analyzed by LC-MS/MS. 293T without expression of FLAG-DDX3X was used as control. Statistical analysis was done in Perseus (version: 1.5.2.6) based on biologically independent samples (n = 3), and potential interactors were determined using Student's two-tailed t-test and visualized by a volcano plot. Red dashed-lines correspond to a given FDR and S0 (FDR = 0.03, S0 = 1.5). Defined interactors were annotated with published SG proteome data [37,40]. The interactors with low number of unique peptides (= 1) or score (< 10) were not annotated in this process. See also **Supplementary Datasheet 2**. (**b**) Comparison of the DDX3X WT interactome with those of acetyl-mimic/dead mutants (n = 3 biologically independent samples). Different interactors were determined using Student's two-tailed t-test and visualized by a volcano plot. Red dashed-lines correspond to a given FDR and S0 (FDR = 0.03, S0 = 1.5).

Supplementary Fig. 26

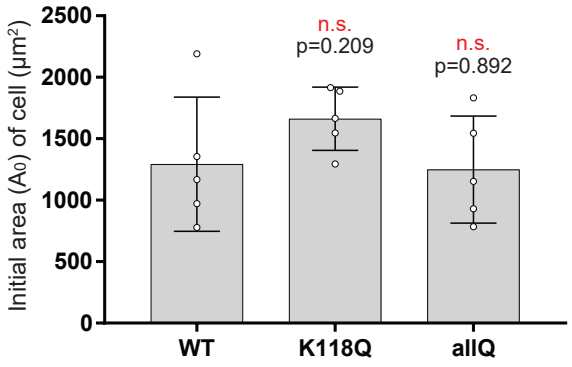
a

	T (min)	k (min^{-1})	r (X 10^{-3})
WT	13.2 ± 2.3 (-)	0.11 ± 0.014 (-)	10.1 ± 3.3 (-)
K118Q	13.5 ± 3.9 (p=0.898, n.s.)	0.12 ± 0.061 (p=0.797, n.s.)	4.0 ± 1.6 (p=0.010, **)
allQ	16.4 ± 3.1 (p=0.140, n.s.)	0.55 ± 0.41 (p=0.071, n.s.)	1.4 ± 0.6 (p=0.0009, ***)

b



c

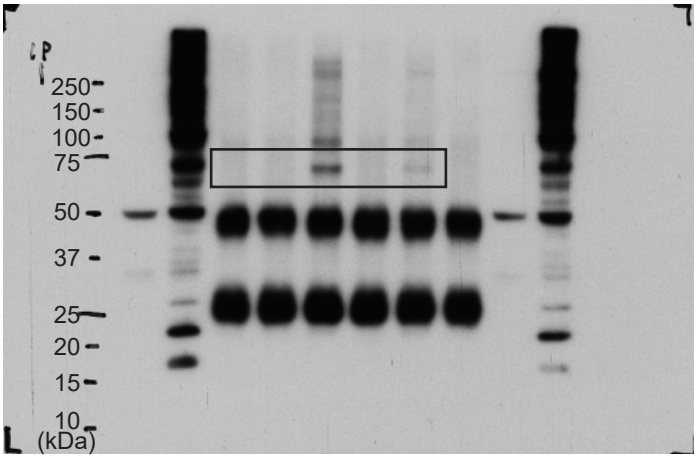


Supplementary Figure 26. Time-dependent growth of mCherry-DDX3X SGs (related to Figure 6).

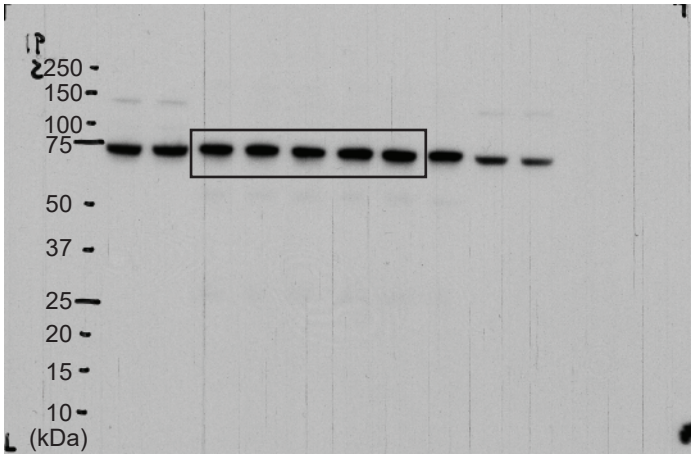
(**a**) The time point at which cells were treated with arsenite was set as t = 0, and the initiation time of SGs (T) was defined as the time point at which SGs were observed for the first time. The obtained total SG area was regressed on Avrami equation and two parameters A_F and k were estimated from this fit. The initial area A_0 was measured at t = 0 for each cell, and the scaling ratio r (= A_F / A_0) was obtained. The best estimate (mean) and associated error (SD) for k and r were calculated from the fit (n = 5 cells for each construct). P values were determined by Student's two-tailed t-test; (*) P < 0.05, (**) P < 0.01 and (***) P < 0.001. (**b**) WT MEFs transiently expressing mCherry-DDX3X WT, K118Q or allQ mutant were treated with 0.5 mM arsenite, and time-lapse images were captured. Total area of SGs within a single cell was measured with ImageJ software, and plotted (n = 5 cells for each construct) after the time point at which cells were treated with arsenite (t = 0). (**c**) Quantification of initial cell area A_0, showing that there is no significant difference between the samples. Mean and SD values are presented (n = 5 cells for each construct). P values were determined by Student's two-tailed t-test; P > 0.05.

Supplementary Fig. 27_1

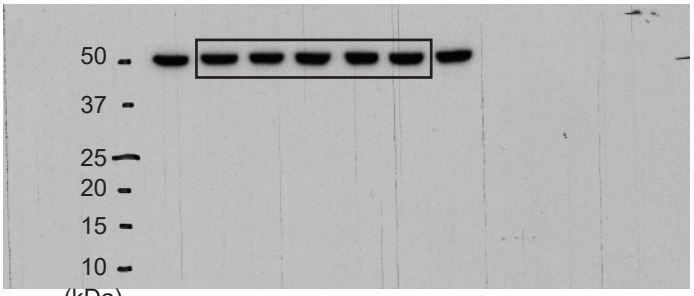
a Fig.1d



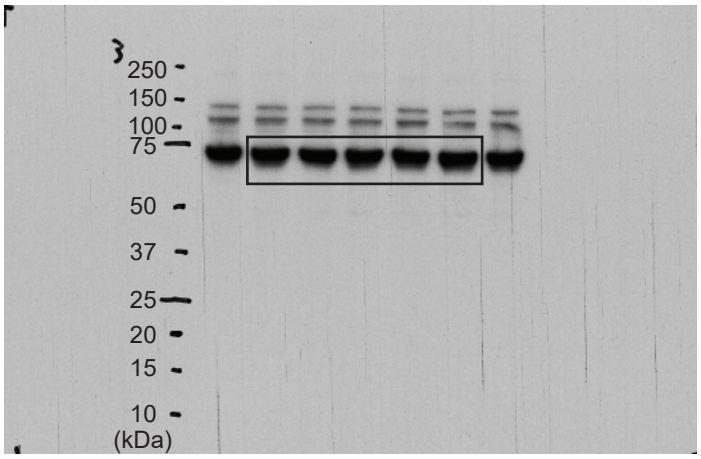
IP: DDX3X
WB: AcK



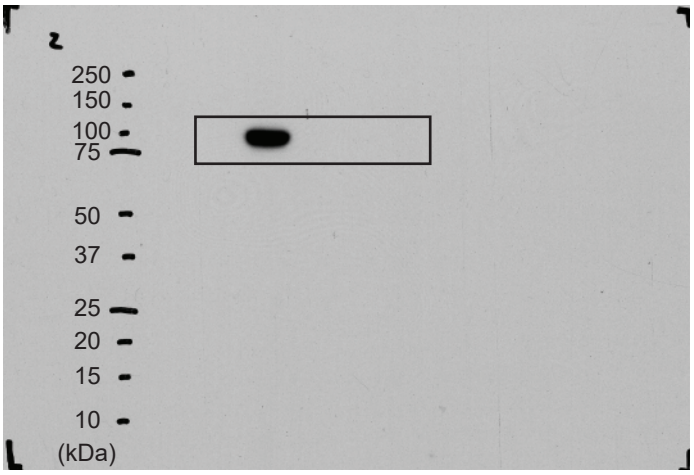
IP: DDX3X
WB: DDX3X



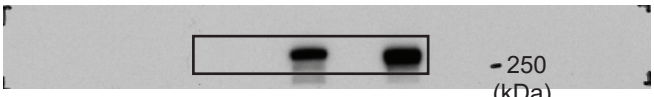
WB: α-tubulin



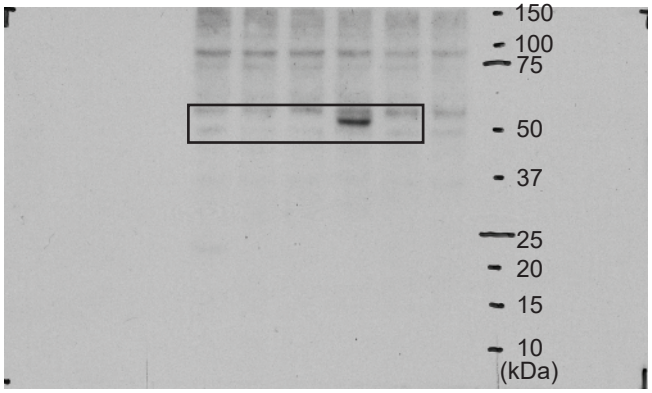
WB: DDX3X



WB: FLAG
(for FLAG-PCAF)



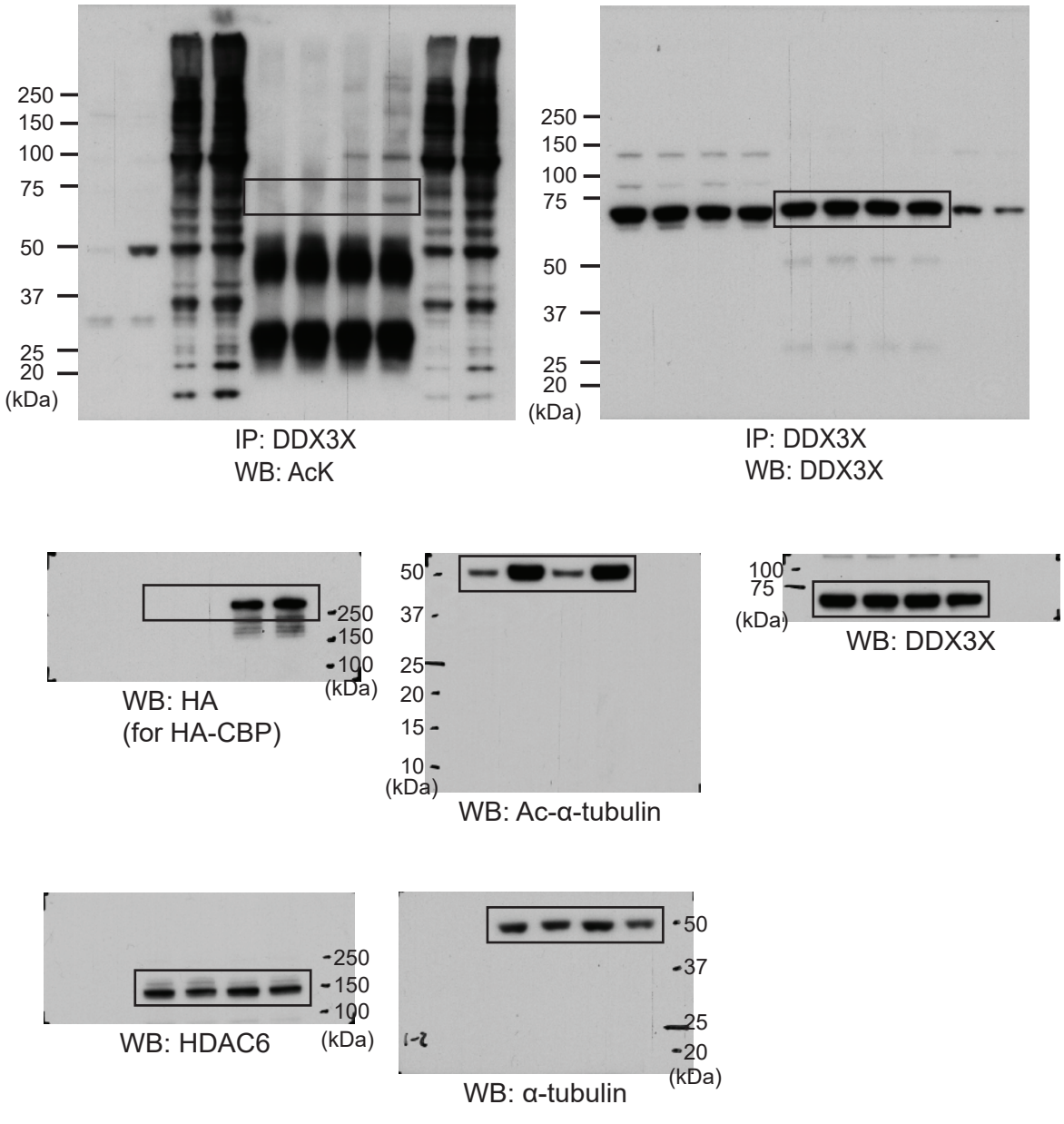
WB: HA
(for HA-CBP/HA-p300)



WB: Myc
(for Myc-Tip60)

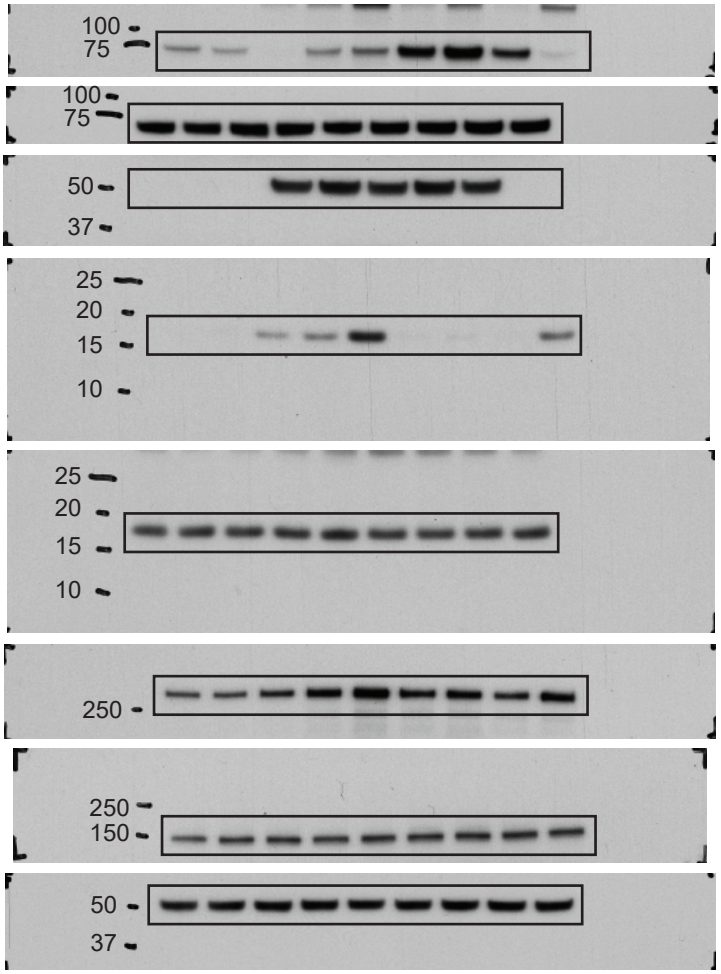
Supplementary Fig. 27_2

b Fig.1e



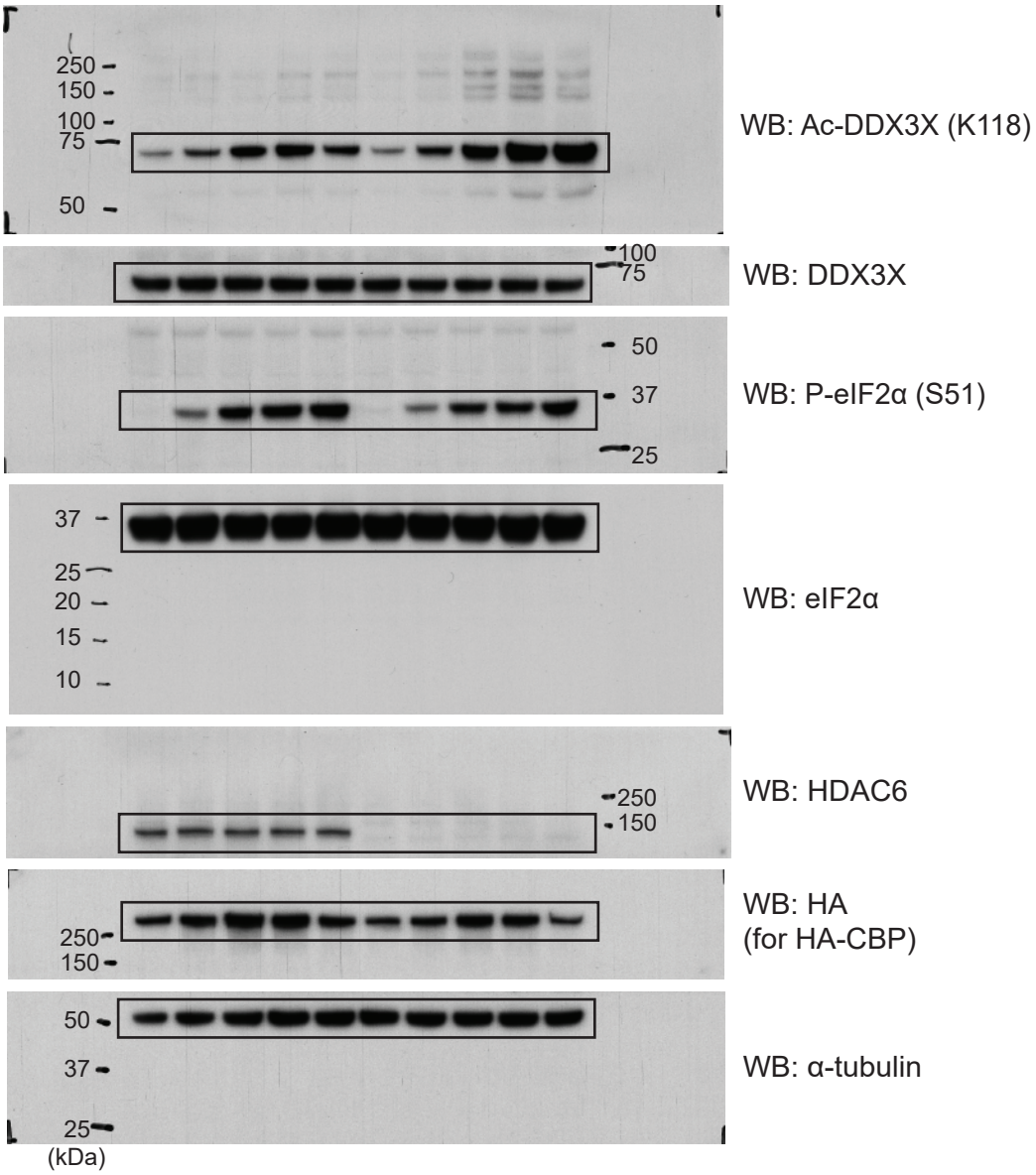
Supplementary Fig. 27_3

c Fig.1f



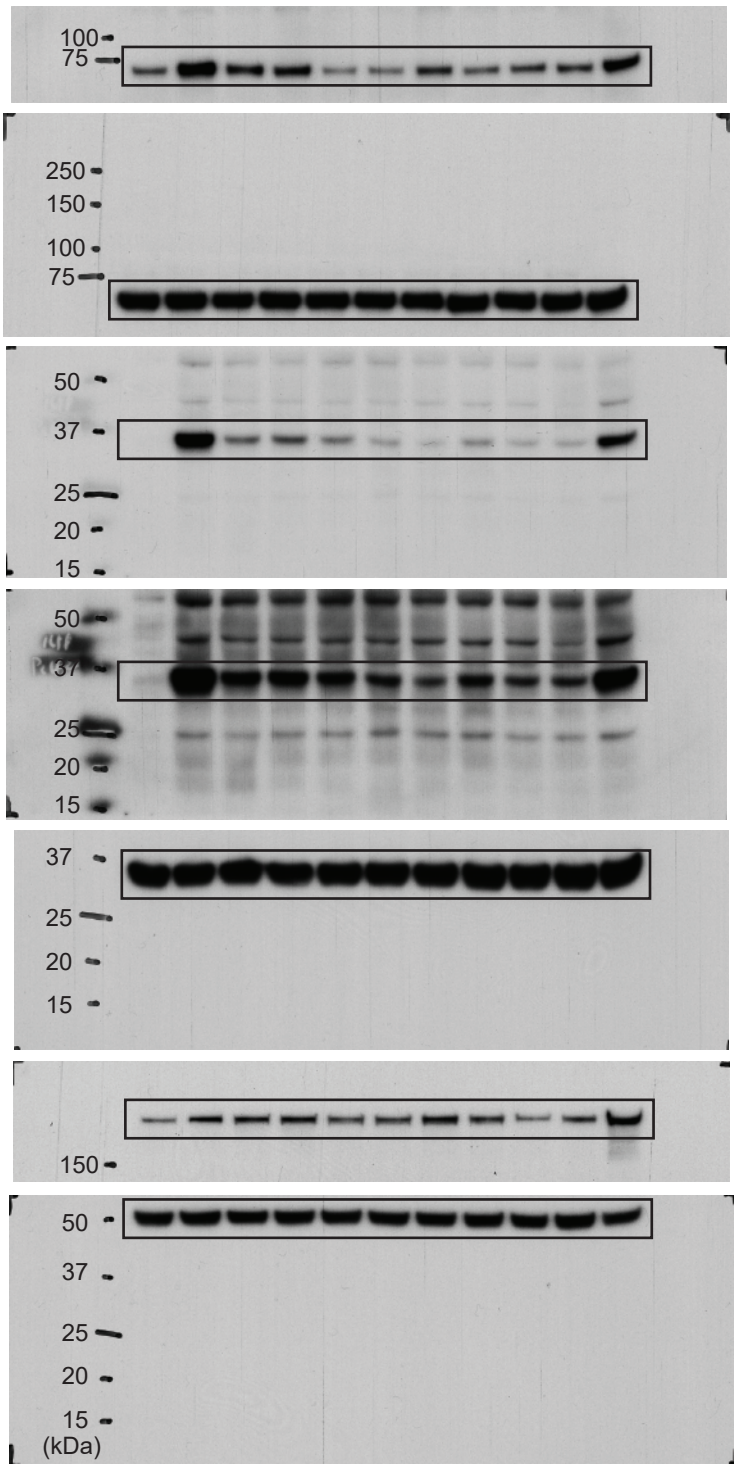
Supplementary Fig. 27_4

d Fig.2a



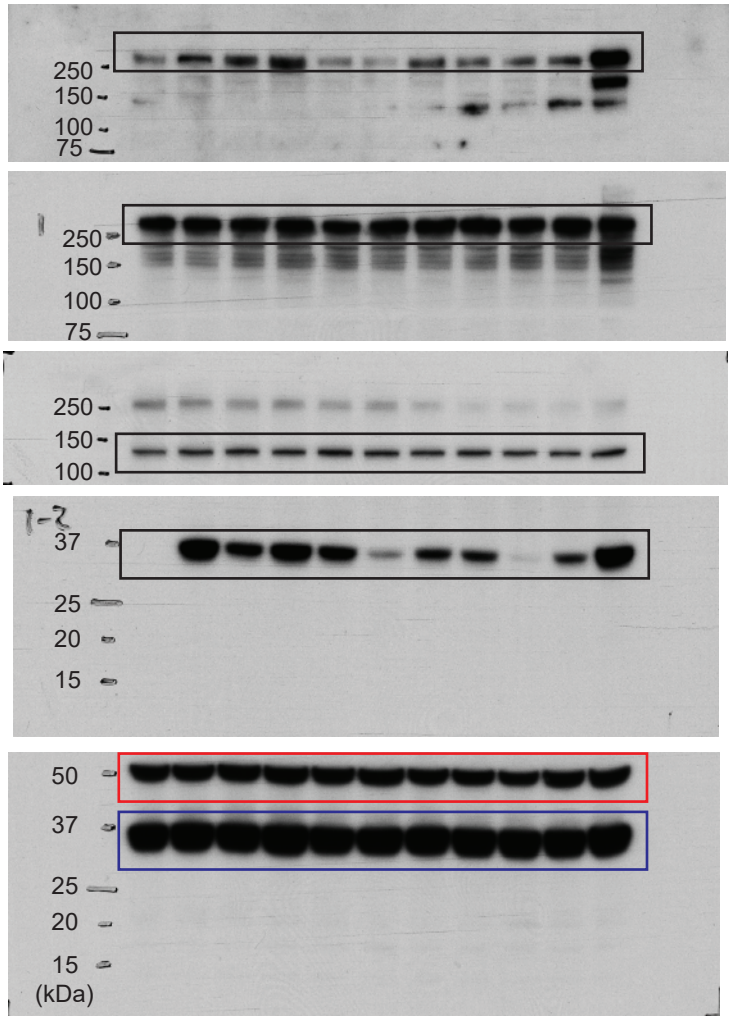
Supplementary Fig. 27_5

e Fig.2c



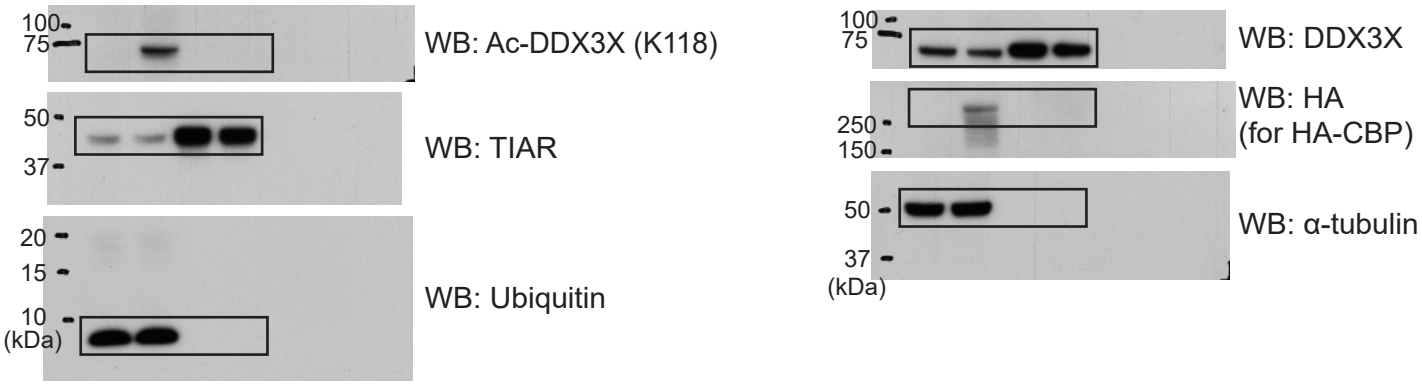
Supplementary Fig. 27_6

f Fig.2e



Supplementary Fig. 27_7

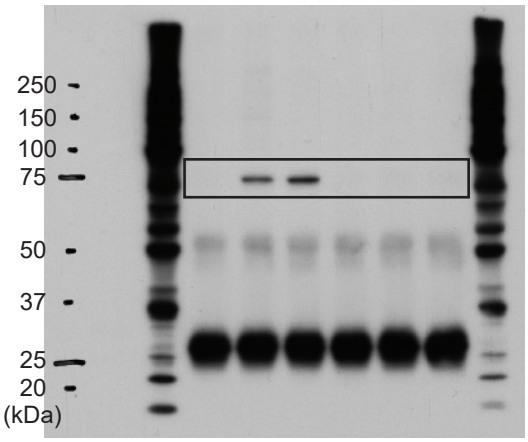
g Fig.3e



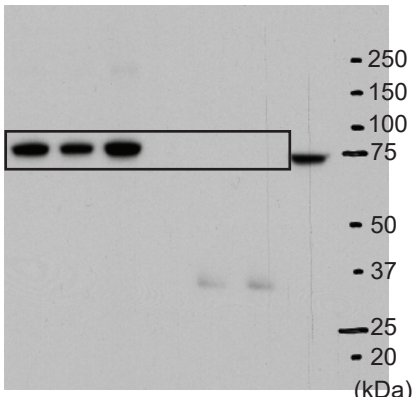
h Fig.4a

Supplementary Fig. 27_8

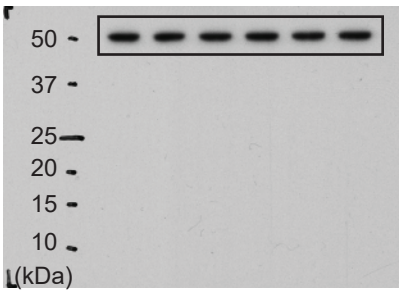
i Supplementary Fig. 5a



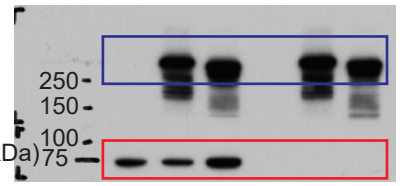
IP: FLAG
WB: AcK



IP: FLAG
WB: FLAG (for FLAG-DDX3X)

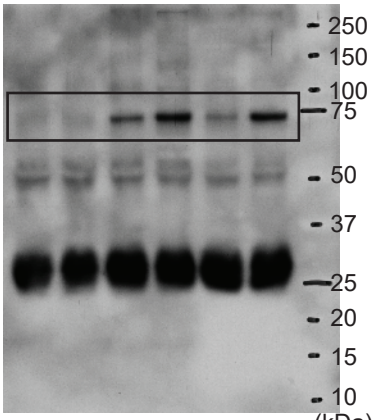


WB: α-tubulin

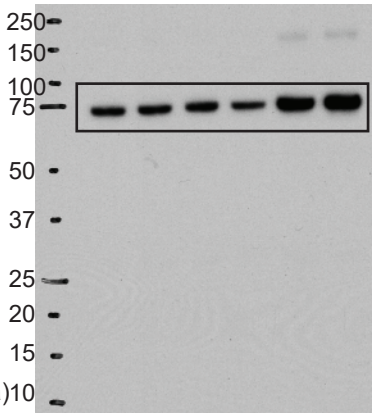


WB: FLAG (for FLAG-DDX3X)
and HA (for HA-CBP/HA-p300)

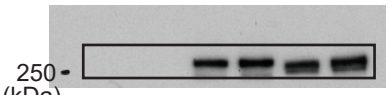
j Supplementary Fig. 5b



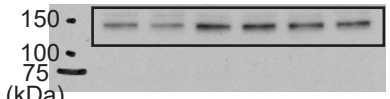
IP: FLAG
WB: AcK



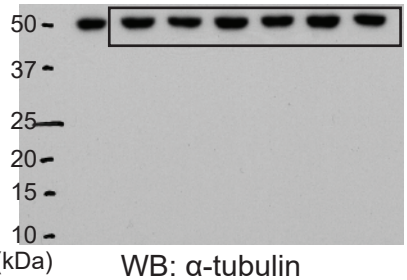
IP: FLAG
WB: FLAG (for FLAG-DDX3X)



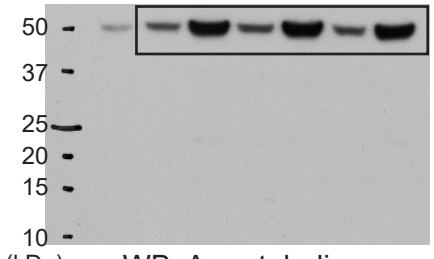
WB: HA
(for HA-CBP/HA-p300)



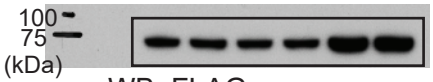
WB: HDAC6



WB: α-tubulin



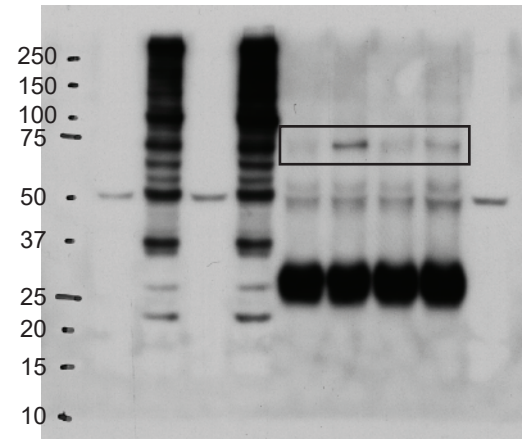
WB: Ac-α-tubulin



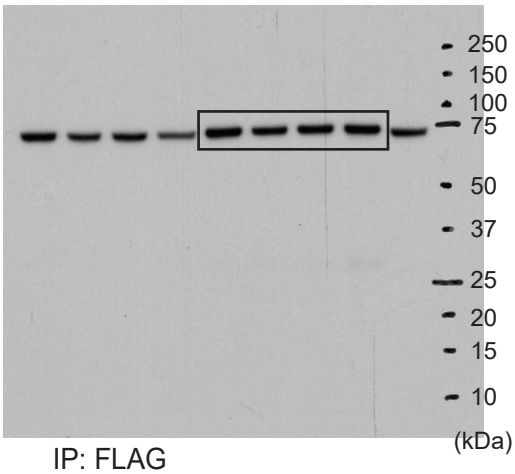
WB: FLAG
(for FLAG-DDX3X)

Supplementary Fig. 27_9

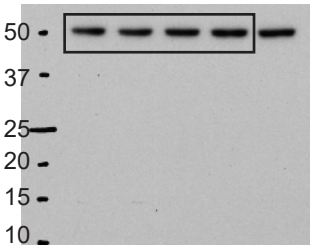
k Supplementary Fig. 5c



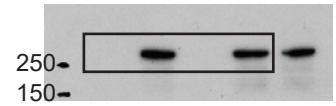
IP: FLAG
WB: AcK



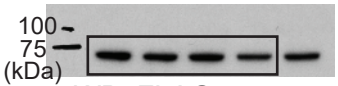
IP: FLAG
WB: FLAG (for FLAG-DDX3X)



WB: α-tubulin

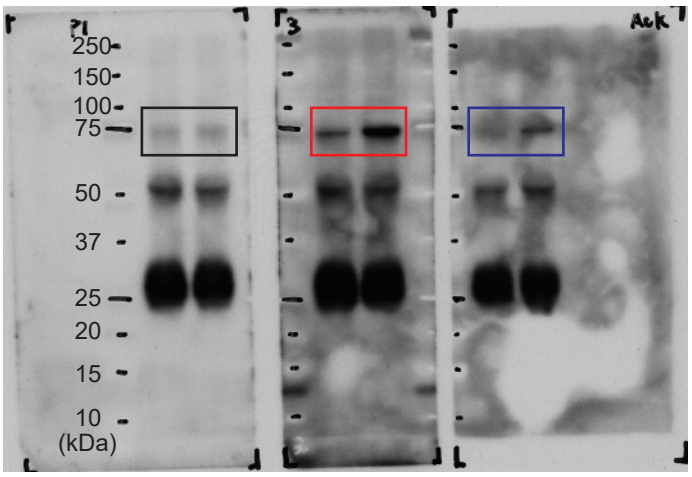


WB: HA
(for HA-CBP)



WB: FLAG
(for FLAG-DDX3X)

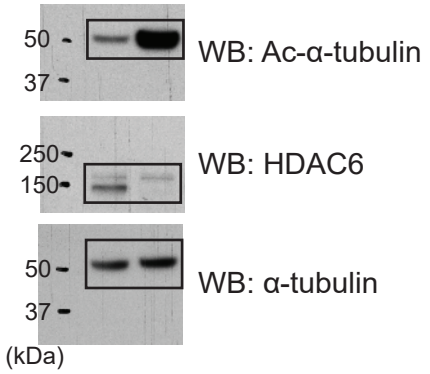
l Supplementary Fig. 7a



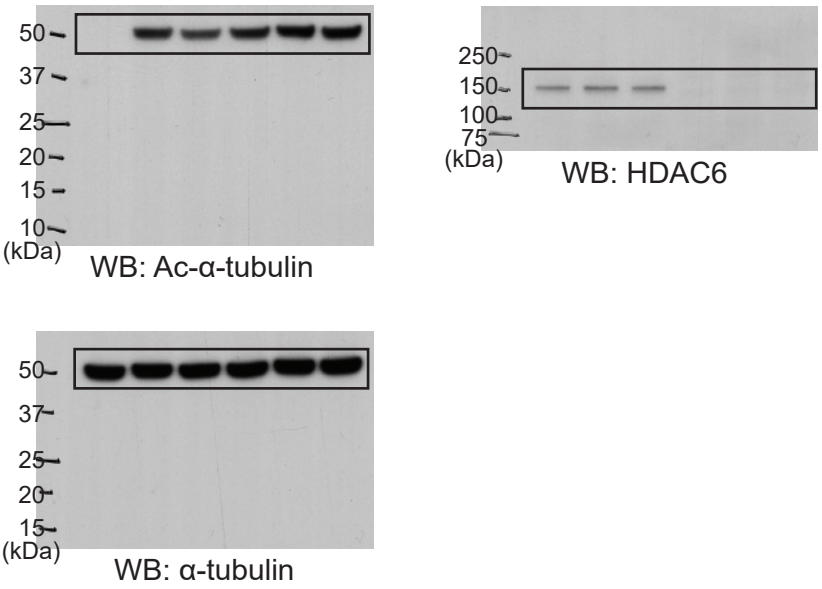
WB: Pre-immune, Immune and AcK

Supplementary Fig. 27_10

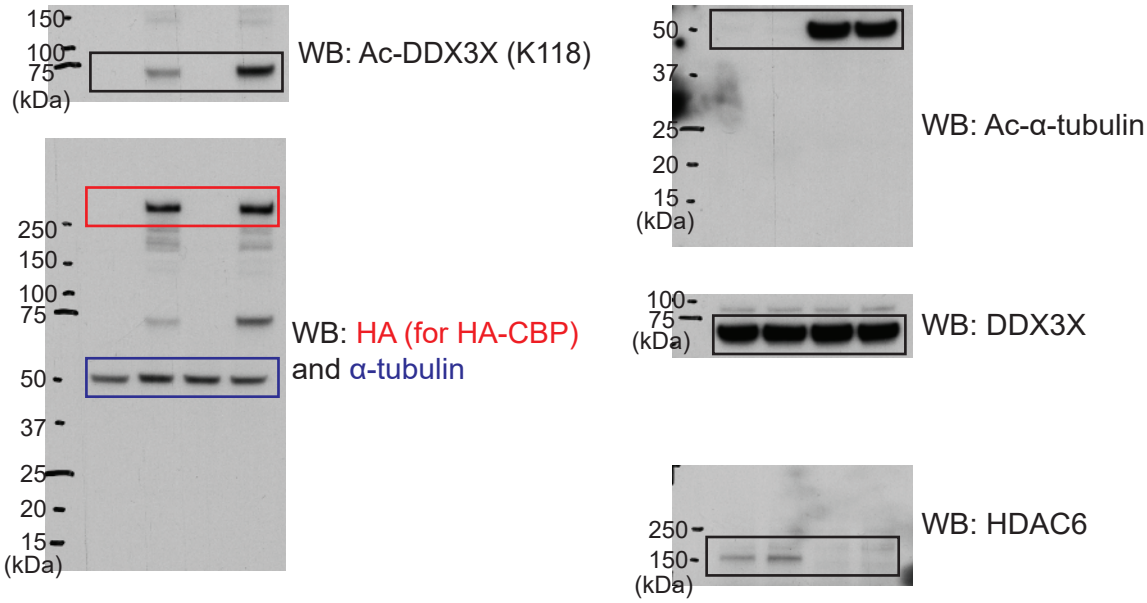
m Supplementary Fig. 8b



n Supplementary Fig. 8c

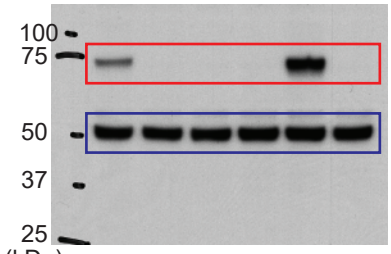


o Supplementary Fig. 8d

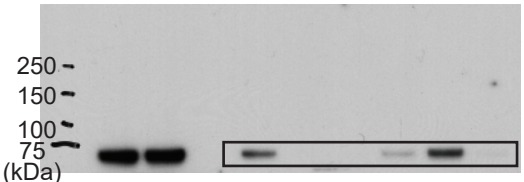


Supplementary Fig. 27_11

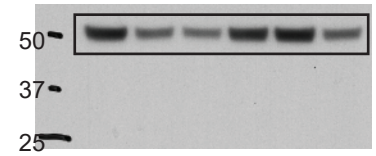
p Supplementary Fig. 9a



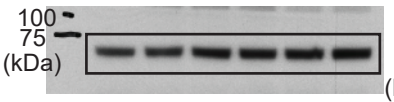
WB: Ac-DDX3X (K118)
and α-tubulin



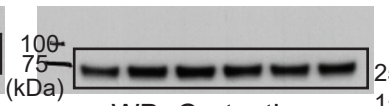
WB: Ac-Cortactin



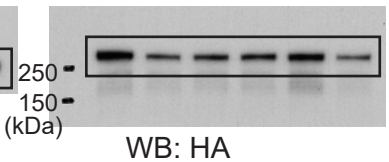
WB: Ac-α-tubulin



WB: DDX3X



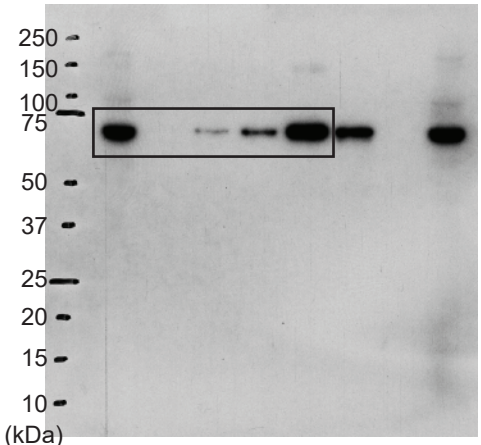
WB: Cortactin



WB: HA
(for HA-CBP)

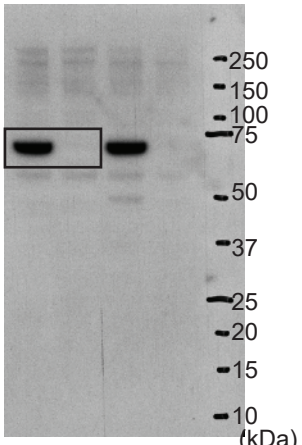
WB: HDAC6

q Supplementary Fig. 13b

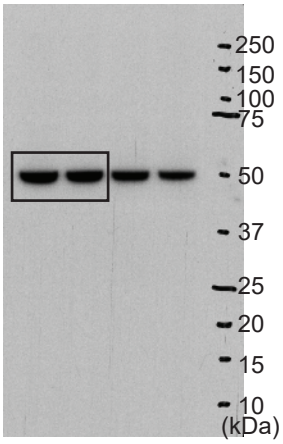


WB: DDX3X

r Supplementary Fig. 16b



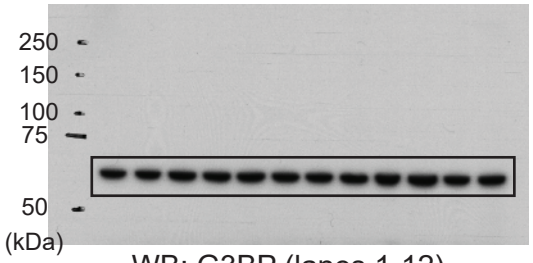
WB: DDX3X



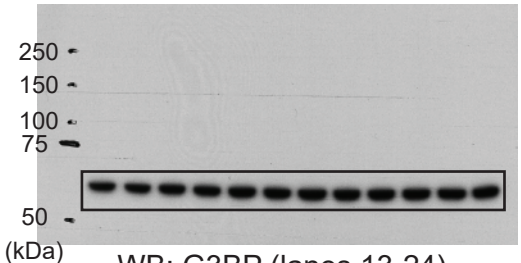
WB: α-tubulin

Supplementary Fig. 27_12

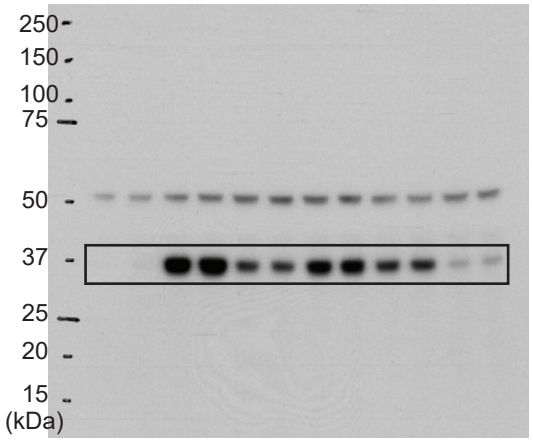
s Supplementary Fig. 17c



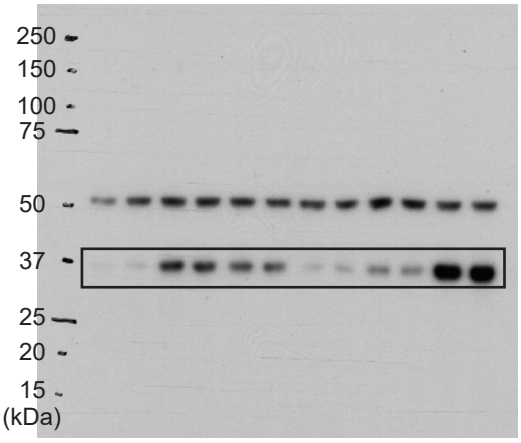
WB: G3BP (lanes 1-12)



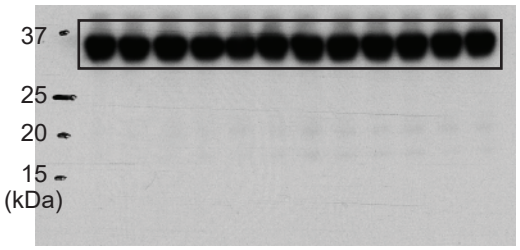
WB: G3BP (lanes 13-24)



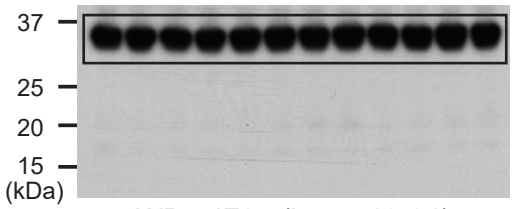
WB: P-eIF2α (S51) (lanes 1-12)



WB: P-eIF2α (S51) (lanes 13-24)

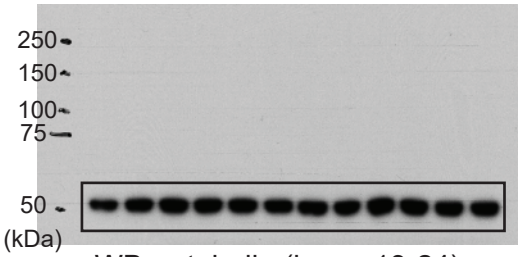


WB: eIF2α (lanes 1-12)



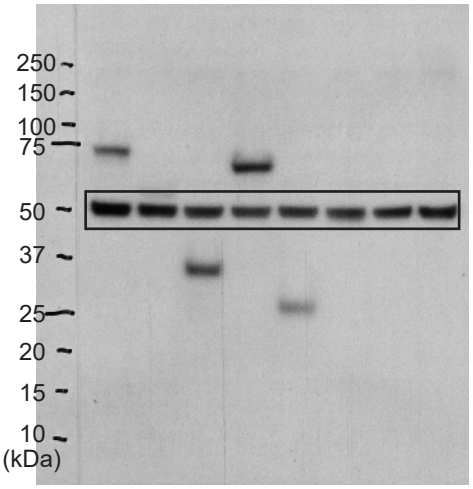
WB: eIF2α (lanes 13-24)

WB: α-tubulin (lanes 1-12)



WB: α-tubulin (lanes 13-24)

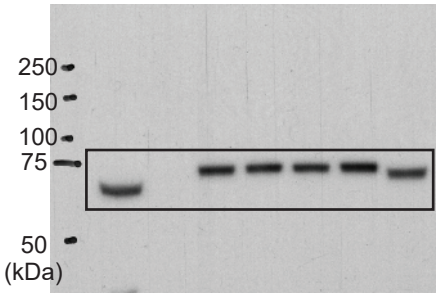
t Supplementary Fig. 18b



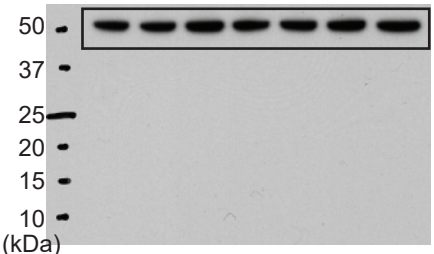
WB: α-tubulin

Supplementary Fig. 27_13

u Supplementary Fig. 20a

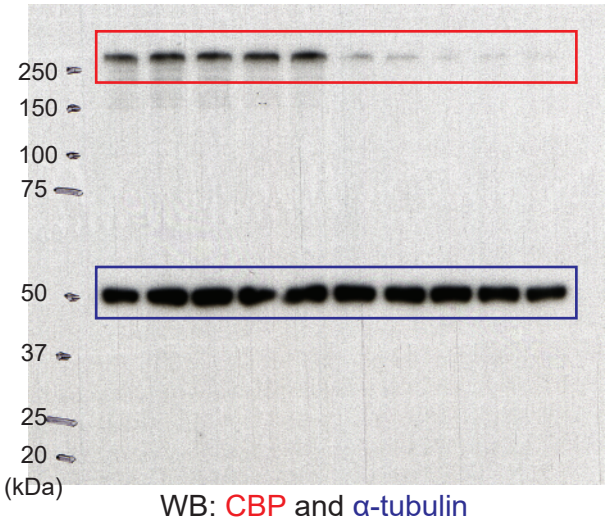


WB: DDX3X

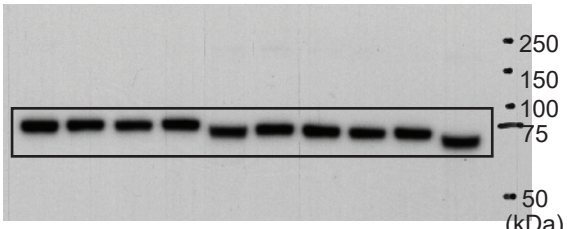


WB: α-tubulin

v Supplementary Fig. 23a



WB: CBP and α-tubulin



WB: DDX3X

Supplementary Figure 27. Original full images of immunoblots.

Boxes indicate cropped images used in the figures. It is important to note that reprobing was avoided for these experiments in order to accurately detect and quantify the acetylation status of DDX3X (**a**, **b**, **i**, **j** and **k**) and CBP (**f**); immunoprecipitated materials were carefully divided and separately loaded on two gels, followed by immunoblotting with the antibodies for target proteins and their acetylated forms (AcK or Ac-CBP (K1535)). Membranes with transferred proteins were cut into several pieces based on the position of molecular-weight size markers, so that several target proteins including loading control could be probed simultaneously (i.e. without exposing the membrane to different antibodies several times). Immunoblotting was repeated with biologically independent samples (n = 3), and representative results are shown for each experiment. (**a**) **Fig. 1d**. (**b**) **Fig. 1e**. (**c**) **Fig. 1f**. (**d**) **Fig. 2a**. (**e**) **Fig. 2c**. (**f**) **Fig. 2e**. (**g**) **Fig. 3e**. (**h**) **Fig. 4a**. (**i**) **Supplementary Fig. 5a**. (**j**) **Supplementary Fig. 5b**. (**k**) **Supplementary Fig. 5c**. (**l**) **Supplementary Fig. 7a**. (**m**) **Supplementary Fig. 8b**. (**n**) **Supplementary Fig. 8c**. (**o**) **Supplementary Fig. 8d**. (**p**) **Supplementary Fig. 9a**. (**q**) **Supplementary Fig. 13b**. (**r**) **Supplementary Fig. 16b**. (**s**) **Supplementary Fig. 17c**. (**t**) **Supplementary Fig. 18b**. (**u**) **Supplementary Fig. 20a**. (**v**) **Supplementary Fig. 23a**.

Supplementary Notes

Supplementary Datasheet 1: Total DDX3X SG volume (related to Figure 5).
Summary of total DDX3X SG volume under different conditions tested in this study (n = 50 cells in total, from biologically independent experiments).

Supplementary Datasheet 2: DDX3X-interactome (related to Figure 6).
Interactomes of each DDX3X WT and acetyl-mimic/dead mutant. Statistical analysis was done in Perseus (version: 1.5.2.6) based on biologically independent samples (n = 3), and potential interactors were determined using Student's two-tailed t-test (FDR = 0.03, S0 = 1.5). Defined interactors were annotated with published SG proteome data [37,40]. The interactors with low number of unique peptides (= 1) or score (< 10) were not annotated in this process. See also **Supplementary Fig. 25**.

Supplementary Datasheet 3: Raw data used for mathematical modelling of SG growth (related to Figure 6).
WT MEFs transiently expressing mCherry-DDX3X WT, K118Q or allQ mutant were treated with 0.5 mM arsenite, and time-lapse images were captured. Total area of SGs within a single cell was measured with ImageJ software (arsenite addition was at t = 0). The initiation time of SGs (T) was defined as the time point at which SGs where observed for the first time. The obtained total SG area was regressed on Avrami equation and two parameters A_F and k were estimated from this fit. The initial area A_0 was measured at t = 0 for each cell, and the scaling ratio r (= A_F / A_0) was obtained. The best estimate (mean) and associated error (SD) for k and r were calculated from the fit (n = 5 cells for each construct). See also **Fig. 6a** and **Supplementary Fig. 26**.

Supplementary Movie 1. Fusion behavior of DDX3X-IDR1 droplet (related to Figure 3).
DDX3X-IDR1 (250 μM, spiked with Alexa Fluor 488 - labelled IDR1 at a molar ratio of 100:1) in 20 mM Tris, pH 7.5, 62.5 mM NaCl, 2 mM DTT forms droplets at R.T. Time-lapse images at the bottom of sample chamber were captured by spinning-disk confocal

microscopy (Frame interval: 1 sec) over 300 frames and edited as a movie (30 fps). Representative result is shown (n = 5 independent experiments). Scale bar, 10 μm.

Supplementary Movie 2. DDX3X-IDR1 droplet disappearance by LLPS following a temperature increase (related to Figure 3).

DDX3X-IDR1 (150 μM, spiked with Alexa Fluor 488 - labelled IDR1 at a molar ratio of 100:1) was dissolved in 20 mM Tris, pH 7.5, 150 mM NaCl, 2 mM DTT. The sample temperature was increased from 5°C to 30°C and decreased to 5°C, leading to dissolution and condensation of droplets. During this, time-lapse images were captured by spinning-disk confocal microscopy (Frame interval: 10 sec) over 41 frames and edited as a movie (5 fps). Representative result is shown (n = 6 independent experiments). Scale bar, 10 μm. Quantification of droplet number related to this movie is in **Supplementary Fig. 11d**.

Supplementary Movie 3. Liquid-like properties of mCherry-DDX3X SGs (related to Figure 4).

WT MEFs transiently expressing mCherry-DDX3X were treated with 0.5 mM arsenite, and time-lapse images were captured by spinning-disk confocal microscopy (13 z-stack for each frame, frame interval: 30 sec) over 120 frames and edited as a movie (10 fps). Representative result is shown (n = 5 biologically independent experiments). Scale bar, 10 μm.

Supplementary Movie 4. SG formation of WT DDX3X and its mutants (related to Figure 6).

WT MEFs transiently expressing each mCherry-DDX3X WT, K118Q and allQ mutant were treated with 0.5 mM arsenite, and time-lapse images were captured by spinning-disk confocal microscopy (13 z-stack for each frame, frame interval: 30 sec) over 120 frames and edited as a movie (10 fps). Representative results are shown (n = 5 cells for each construct). Left: WT, middle: K118Q and right: allQ. Scale bar: 10 μm.

Supplementary Information References:

58 Szklarczyk, D. *et al.* The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic acids research* **45**, D362-D368, doi:10.1093/nar/gkw937 (2017).

59 Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J. & Gevaert, K. Improved visualization of protein consensus sequences by iceLogo. *Nature methods* **6**, 786-787, doi:10.1038/nmeth1109-786 (2009).

2.2 Structural insights into HDAC6 tubulin deacetylation and its selective inhibition

(Published, Nature Chemical Biology)

ARTICLE
PUBLISHED ONLINE: 25 JULY 2016 | DOI: 10.1038/NCHEMBIO.2140

Structural insights into HDAC6 tubulin deacetylation and its selective inhibition

Yasuyuki Miyake[1,4], Jeremy J Keusch[1,4], Longlong Wang[1], Makoto Saito[1], Daniel Hess[1], Xiaoning Wang[2], Bruce J Melancon[2], Paul Helquist[2], Heinz Gut[1] & Patrick Matthias[1,3]*

We report crystal structures of zebrafish histone deacetylase 6 (HDAC6) catalytic domains in tandem or as single domains in complex with the (*R*) and (*S*) enantiomers of trichostatin A (TSA) or with the HDAC6-specific inhibitor nexturastat A. The tandem domains formed, together with the inter-domain linker, an ellipsoid-shaped complex with pseudo-twofold symmetry. We identified important active site differences between both catalytic domains and revealed the binding mode of HDAC6 selective inhibitors. HDAC inhibition assays with (*R*)- and (*S*)-TSA showed that (*R*)-TSA was a broad-range inhibitor, whereas (*S*)-TSA had moderate selectivity for HDAC6. We identified a uniquely positioned α-helix and a flexible tryptophan residue in the loop joining α-helices H20 to H21 as critical for deacetylation of the physiologic substrate tubulin. Using single-molecule measurements and biochemical assays we demonstrated that HDAC6 catalytic domain 2 deacetylated α-tubulin lysine 40 in the lumen of microtubules, but that its preferred substrate was unpolymerized tubulin.

Acetylation on lysine residues of histone and other proteins has been recognized as a major post-translational modification that affects multiple aspects of protein function. Protein acetylation levels are regulated by the balance of enzymes with opposing activities: histone acetyltransferases (HATs) and HDACs[1]. HDAC6 is the major deacetylase for tubulin, and it also deacetylates proteins such as HSP90 and cortactin, among others[2–4]. It is mostly cytoplasmic with unique characteristics that set it apart from other deacetylases: tandem catalytic domains with the capacity to deacetylate tubulin and the presence of a zinc finger domain with homology to ubiquitin-specific proteases (ZnF-UBP), which binds unanchored ubiquitin. HDAC6 is a central modulator of stress responses and autophagic clearance, essential for the formation of aggresomes or stress granules[5–7]. It also has an important role for regulatory T cells[8], influenza virus infection[9] and in pathological conditions such as cancer, inflammation and neurodegeneration[10,11].

Microtubules (MTs) are assembled from α- and β-tubulin heterodimers to form dynamic cytoplasmic filaments, involved in multiple cellular functions comprising cell cycle, cell shape, cellular motility and intracellular transport of cargos such as vesicles or viruses[12]. MTs are heavily decorated by post-translational modifications including acetylation, glutamylation, tyrosination or phosphorylation, which have been proposed to regulate their properties, stability and functions[13,14]. α-TAT is the only acetyltransferase targeting Lys40, a residue located in a flexible loop of α-tubulin in the luminal side of MTs[15,16]. α-TAT prefers MTs over α/β-tubulin heterodimers for the efficient acetylation of α-tubulin Lys40 (refs. 17,18), and stochastic acetylation of MTs by α-TAT had been recently demonstrated[19]. Deacetylation of tubulin is promoted by HDAC6 (refs. 2,20) and by the NAD-dependent class III deacetylase SIRT2 (ref. 21). HDAC6 and SIRT2 interact and may function together[21,22]. However, alteration of HDAC6 levels is sufficient to increase tubulin acetylation, and fibroblasts lacking HDAC6 have fully acetylated tubulin[23]. HDAC6 also interacts with plus-end tip-binding proteins such as EB1 or Arp1 (ref. 24), indicating that it might deacetylate the end of microtubules. It is not firmly established which is the preferred substrate of HDAC6, α/β-tubulin dimers or polymeric MTs[2,25]. It has recently been reported that interaction between HDAC6 and tubulin is direct[26] but also that septins facilitate interaction between HDAC6 and acetylated α-tubulin[27].

HDACs comprise 11 family members and are considered promising targets in a number of pathologies, with cancer being the most advanced indication[28]. In most cases however, the critical HDAC(s) have not been conclusively identified, and the four inhibitors (vorinostat, romidepsin, belinostat and panobinostat) approved for cancer treatment until now all target multiple HDACs[29]. HDAC6 is currently evaluated as a potential therapeutic target in particular in multiple myeloma. HDAC6-selective inhibitors have been developed and clinical trials are underway with a recent HDAC6-selective inhibitor, ricolinostat (ACY-1215)[30].

We solved the crystal structure of both catalytic domains of HDAC6, which together with the inter-domain linker form an ellipsoid-shaped complex with pseudo-twofold symmetry. We structurally and functionally defined features that are critical for HDAC6 to deacetylate its physiologic substrate tubulin, and we showed that HDAC6 prefers tubulin dimers as substrate but that it can stochastically deacetylate MTs. We also determined the crystal structures of individual catalytic domains bound to either enantiomer of TSA or to the HDAC6-specific inhibitor nexturastat A (NextA), and found that (*S*)-TSA had moderate selectivity for HDAC6.

RESULTS

Organization of the HDAC6 tandem catalytic domains

To understand how HDAC6 deacetylates tubulin and other substrates we characterized this multidomain protein by X-ray crystallography. As we did not obtain crystals with the mouse protein, we used the zebrafish ortholog, which efficiently deacetylates tubulin from various sources (**Supplementary Results**, **Supplementary Fig. 1**). To facilitate the comparison between zebrafish and other species, we aligned sequences of HDAC6 proteins (**Supplementary Fig. 2**). We first used proteins containing the tandem catalytic domains to characterize their activity in an HDAC assay using Fluor de Lys as

[1]Friedrich Miescher Institute for Biomedical Research (FMI), Basel, Switzerland. [2]Department of Chemistry & Biochemistry, University of Notre Dame, Notre Dame, Indiana, USA. [3]Faculty of Sciences, University of Basel, Basel, Switzerland. [4]These authors contributed equally to this work.
*e-mail: patrick.matthias@fmi.ch

substrate. We compared wild-type catalytic domains 1 and 2 (CD1-CD2) to proteins with an inactivating substitution in either catalytic domain ($CD1_{H193A}$-CD2 or CD1-$CD2_{H574A}$) or in both domains ($CD1_{H193A}$-$CD2_{H574A}$)[20,31,32]. SDS-PAGE and size-exclusion chromatography with multi-angle light scattering analysis showed that the proteins used were highly pure and homogeneous (**Supplementary Fig. 3a**,**b**). Whereas the wild-type CD1-CD2 fusion was highly active, substitution of the first catalytic site ($CD1_{H193A}$-CD2) had almost no impact. In contrast, substitution of the second site (CD1-$CD2_{H574A}$) strongly impaired the protein but without fully inactivating it, and the doubly substituted variant ($CD1_{H193A}$-$CD2_{H574A}$) was fully inactive (**Supplementary Fig. 3c**). This indicates that the first catalytic domain of zebrafish HDAC6 has a weak but measurable enzymatic activity (see below). When tested in our assay, the full-length human HDAC6 protein showed weaker activity (**Supplementary Fig. 3c**).

We then determined structures of the tandem catalytic domains in complex with NextA (CD1-CD2, residues 25–831, 2.9 Å, interdomain linker present; and residues 40–831, 2.0 Å, interdomain linker cleaved), CD1 in complex with TSA (CD1-TSA, residues 40–418, 1.5 Å), CD2-TSA (residues 441–831, 1.6 Å), as well as of the ZnF-UBP domain (residues 974–1081, 1.9 Å; highly similar to the human ortholog[33]; **Fig. 1a** and **Supplementary Fig. 4**). Data collection and refinement statistics are summarized in **Supplementary Table 1**. In the CD1-CD2 crystal structure both CD1 and CD2 adopted a classical arginase-deacetylase fold[34,35] and had a conserved deacetylase active site ~50 Å apart. The two domains were closely attached to each other (**Fig. 1b** and **Supplementary Fig. 5**). The large domain-domain interface was formed by helices H13, H14, H15 and H18 of CD1, and H32, H33 and H34 of CD2; by loops connecting helices H17 and H18 of CD1, and H36 and H37 of CD2; as well as by the linker (418–442) connecting the two domains and by the C-terminal part of CD2 (794–806). This resulted in a large buried surface area of ~2,100 Å^2 on each domain upon formation of the complex. CD1 and CD2 were structurally very similar (r.m.s. deviation = 1.0 Å, 354 C_α atoms, 45% sequence identity), and the same structural elements were engaged in the domain-domain interface, resulting in a pseudo-twofold axis running along the interface perpendicular to the α-helices on both sides (**Fig. 1b** and **Supplementary Fig. 5**).

We next determined structures of individual CD1 and CD2 domains in complex with the pan-HDAC inhibitor TSA. The isolated CD1 domain contained (*R*)-TSA in its catalytic center (**Fig. 1c**), and we obtained the structure of CD2-TSA as we were trying to crystallize the tandem domains with TSA: during crystallization, proteolysis repeatedly took place and liberated the individual domains, yielding crystals of CD2 with (*S*)-TSA bound (**Fig. 1d**). Backbones of CD1 and CD2 were highly similar also in single domain structures (r.m.s. deviation = 0.77 Å). Superposition of other HDAC structures revealed variations in the N-terminal and C-terminal part of the HDAC6 catalytic domains (**Supplementary Fig. 6**). A notable feature is the presence of a uniquely positioned 10-residue α-helix, referred to as H6 in CD1 and H25 in CD2, found in each catalytic domain near the active sites, as well as a unique loop at the N terminal part of each catalytic domain: H1-H2 in CD1 and H20-H21 in CD2, containing Trp78 and Asp79 (CD1) and Trp459 and Asp460 (CD2) (**Fig. 1c**,**d** and **Supplementary Fig. 2**).

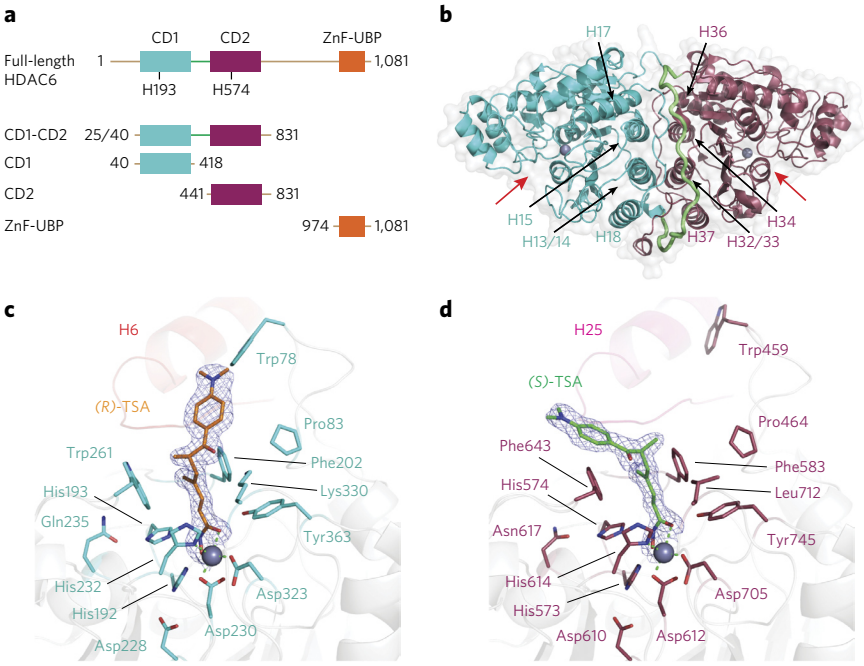


Figure 1 | Overall structure of HDAC6 catalytic domains. (**a**) HDAC6 domain architecture and constructs used for structural studies. This color scheme is used throughout all figures. Amino acid boundaries of all constructs are indicated; for CD1-CD2, two alternative proteins were made starting at position 25 or at position 40. (**b**) Cartoon representation of the CD1-CD2 crystal structure, with CD1 and CD2 in cyan and dark red, respectively, and the interdomain linker in green. Metal ions are shown as gray spheres and red arrows point to the substrate binding clefts. Helices engaged in the CD1-CD2 interface are labeled. (**c**,**d**) Magnified view into inhibitor binding sites of CD1-(*R*)-TSA (**c**) and CD2-(*S*)-TSA (**d**) structures. Residues interacting with TSA or involved in catalysis are drawn as sticks in cyan (CD1) and dark red (CD2); catalytic domain backbones are shown as light gray cartoons with helices H6 and H25 in light red and pink. (*R*)- and (*S*)-TSA are shown as orange (CD1) and green (CD2) sticks with sigma-A-weighted $2mF_o - DF_c$ composite omit electron density maps displayed as blue mesh (1 σ). Zinc ions are shown as gray spheres, and polar interactions involved in Zn^{2+} binding are represented as green dashed lines.

CD1 and CD2 catalytic pockets, and TSA binding

Both HDAC6 CD1 and CD2 active sites were highly conserved and featured the typical narrow hydrophobic channel formed by residues Pro83, Gly201, Phe202 and Trp261 in CD1, and Pro464, Gly582, Phe583, Phe643 and Leu712 in CD2 (**Fig. 1c**,**d**). The Zn^{2+} ion was coordinated by Asp230, His232 and Asp323 in CD1, and Asp612, His614 and Asp705 in CD2. In CD1, the two charge relay systems consisted of the His192-Asp228 and His193-Gln235 dyads, with the latter normally being a His-Asn pair. In contrast, CD2 had the classical dyad arrangement with His573-Asp610 and His574-Asn617. Lastly, Tyr residues located next to the catalytic zinc ion and thought to stabilize the reaction intermediate, as in class I enzymes[35], were conserved in CD1 (Tyr363) and CD2 (Tyr745). A noteworthy difference present in all HDAC6 sequences when comparing CD1 and CD2 is the use of the bulkier Trp261 in CD1, instead of the usual phenylalanine (Phe643 in CD2), to form one wall of the hydrophobic acetylated lysine binding channel (**Supplementary Figs. 2** and **7**).

(*R*)- and (*S*)-TSA binding to CD1 and CD2 were highly similar for the hydroxamate moiety that complexes the corresponding Zn^{2+} ion in a bidentate fashion using its carbonyl and hydroxyl oxygens (**Fig. 1c**,**d**). The unsaturated aliphatic TSA linkers were nearly planar and sandwiched between the aromatic side chains of Phe202

Table 1 | Inhibitory profile of (*R*)-TSA and (*S*)-TSA against zebrafish and human HDAC6 as well as against human HDAC1–11

Isoform	(*R*)-TSA IC_{50} (nM)	(*S*)-TSA IC_{50} (nM)
Zebrafish HDAC6 (CD1-CD2)	5.45 ± 0.62	9.88 ± 1.01
Human HDAC6	4.67 ± 0.06	11.1 ± 0.62
Human HDAC1	5.76 ± 1.05	206.30 ± 15.84
Human HDAC2	17.81 ± 1.08	612.65 ± 116.60
Human HDAC3	8.09 ± 0.28	320.80 ± 27.01
Human HDAC4	9,613 ± 2,329.21	6,341 ± 627.91
Human HDAC5	4,385 ± 1,248.75	6,325 ± 117.38
Human HDAC7	3,499.50 ± 123.74	1,823.50 ± 6.36
Human HDAC8	410.50 ± 43.27	312.20 ± 3.96
Human HDAC9	8,861 ± 60.10	4,824 ± 228.40
Human HDAC10	29.19 ± 0.06	403.35 ± 10.25
Human HDAC11	3,642.50 ± 683.77	2,684.00 ± 398.81

IC_{50} is the mean of two experiments ± s.d. obtained from curve fitting of 10-point enzyme assays with threefold serial dilution. HDAC6 assays started at 2 μM inhibitor; HDAC4, HDAC5 and HDAC9 assays started at 450 μM inhibitor; HDAC1, HDAC2, HDAC3, HDAC7, HDAC8, HDAC10 and HDAC11 assays started at 50 μM inhibitor. Values were extracted from fitting dose-response curves to the data points using GraphPad software.

and Trp261 in CD1, and Phe583 and Phe643 in CD2. Whereas the carbonyl group of the hydroxamate was almost coplanar with the unsaturated aliphatic chain, the best fit to the electron density was achieved by kinking the hydroxylamine toward the zinc ion by ~30°, somewhat resembling the TSA conformation in the HDAC7-TSA complex structure[36] (Protein Data Bank: 3C10). The TSA dimethylamino-phenyl CAP group used the first part of a CD1 groove formed by H6 and loop H1-H2 for hydrophobic interaction with Trp78 and Phe202 side chains. Owing to crystallization of the (*S*) enantiomer in CD2, we did not see such an orientation of the CAP group, and it instead interacted with the Phe643 side chain. The side chains of Trp78 in CD1 and of the corresponding Trp459 in CD2 assumed different conformations in their respective grooves when engaged in ligand binding (CD1) or it was free (CD2), thus pointing to a role in substrate recognition (see below). We present additional differences between CD1 and CD2 and their possible role in substrate recognition in **Supplementary Figure 7**.

(*S*)-TSA vs. (*R*)-TSA inhibition of HDAC6 and other HDACs

It had previously been reported that the unnatural (*S*) enantiomer of TSA is biologically inactive[37] and does not inhibit partially purified HDACs from mouse cells[38]; subsequently most studies have used the natural form, (*R*)-TSA. Our observation that HDAC6 CD2 was bound by (*S*)-TSA in our crystal structure was intriguing, and we set out to characterize the activity of (*S*)-TSA vs. (*R*)-TSA on zebrafish and human HDAC6, as well as on all other human HDACs. With pure preparations of the two enantiomers[39] (**Supplementary Fig. 8**), we found that (*R*)-TSA and (*S*)-TSA inhibited zebrafish and human HDAC6 similarly (half-maximal inhibitory concentration (IC_{50}) for zebrafish was 5.45 nM (*R*-TSA) and 9.88 nM (*S*-TSA), and for human, 4.67 nM (*R*-TSA) and 11.1 nM (*S*-TSA)). In contrast, when we tested other human HDACs, we observed differences between (*R*)-TSA and (*S*)-TSA: class I HDACs HDAC1–3 and class II HDAC10 were all strongly inhibited by (*R*)-TSA but only weakly by (*S*)-TSA, whereas other HDACs, including HDAC6, were inhibited about equally or with only small differences (**Table 1**). Thus, although (*R*)-TSA is a pan-HDAC inhibitor, (*S*)-TSA, the unnatural enantiomer, had *in vitro* moderate selectivity for HDAC6 (~20-fold lower IC_{50} compared to the next closest isoform, HDAC1).

To test whether (*S*)-TSA also has activity *in vivo*, we treated mouse embryonic fibroblasts (MEFs) with (*R*)-TSA or (*S*)-TSA, or with the HDAC6-specific inhibitor NextA. After treatment with inhibitor, we prepared protein lysates and monitored acetylation of tubulin or histone H3 by immunoblotting (**Supplementary Fig. 9**). (*S*)-TSA was active *in vivo*, although weaker than the (*R*) enantiomer. However, acetylation of tubulin increased slightly more rapidly than that of histone H3, indicating a moderate selectivity for HDAC6. As expected, NextA, which *in vitro* has very high selectivity for HDAC6 (600-fold over HDAC1 and >1,000-fold over HDAC2; ref. 40), affected acetylation of tubulin without appreciably impacting histone H3 acetylation.

Conservation of active site and surrounding residues

We wondered whether active site differences between CD1 and CD2 are conserved through evolution, and used ConSurf[41] to build a multiple-sequence alignment to map site-specific conservation scores onto the CD1-CD2 surface (**Supplementary Fig. 10a,b** and **Supplementary Dataset 1**). This analysis indicated that the CD2 active site and substrate recognition region were under more evolutionary pressure for conservation than the CD1 site and may point to different functions for the two domains, with a more important role for CD2 (see below).

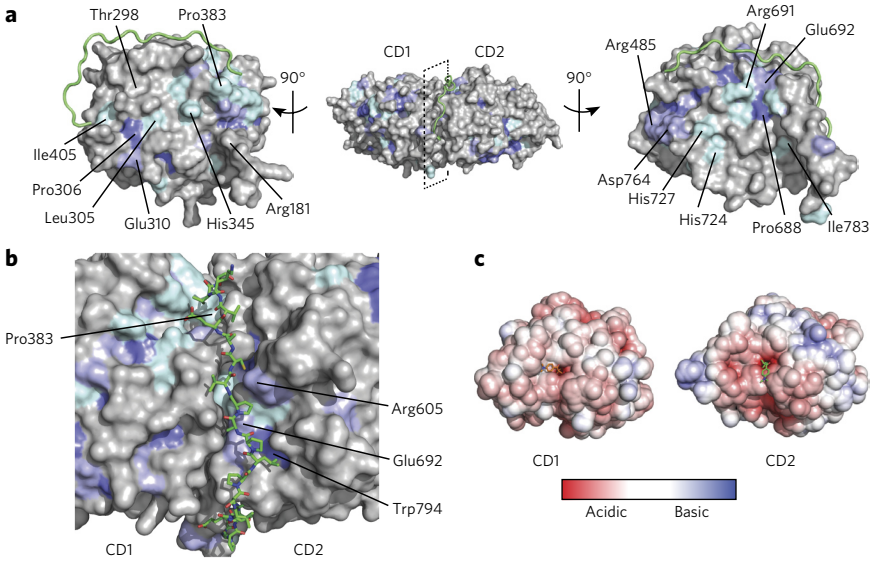


Figure 2 | Conservation analysis of CD1 vs. CD2. (**a**) Site-specific conservation scores computed by ConSurf (http://consurf.tau.ac.il/) from 150 HDAC6 sequences mapped onto the HDAC6 surface. Consurf color grades 9 (dark blue), 8 (light blue) and 7 (cyan) highlight the degree of residue conservation from high to medium. The surface representation of the CD1-CD2 crystal structure is shown in the center. The linker connecting CD1 and CD2 is shown in green as a coil. On the left and right, CD1 and CD2 domains are rotated by 90° around a vertical axis to display the interdomain interface. (**b**) Detailed view on the linker connecting CD1 and CD2, and its interaction with conserved residues on both domains. (**c**) Mapping of the electrostatic surface potential from –8 *kT/e* (red) to +8 *kT/e* (blue) onto the CD1 and CD2 domains (APBS implemented in PyMOL). The domains have been reoriented to show the TSA-bound catalytic sites from the top. Here the linker is included in the surface representation of CD1.

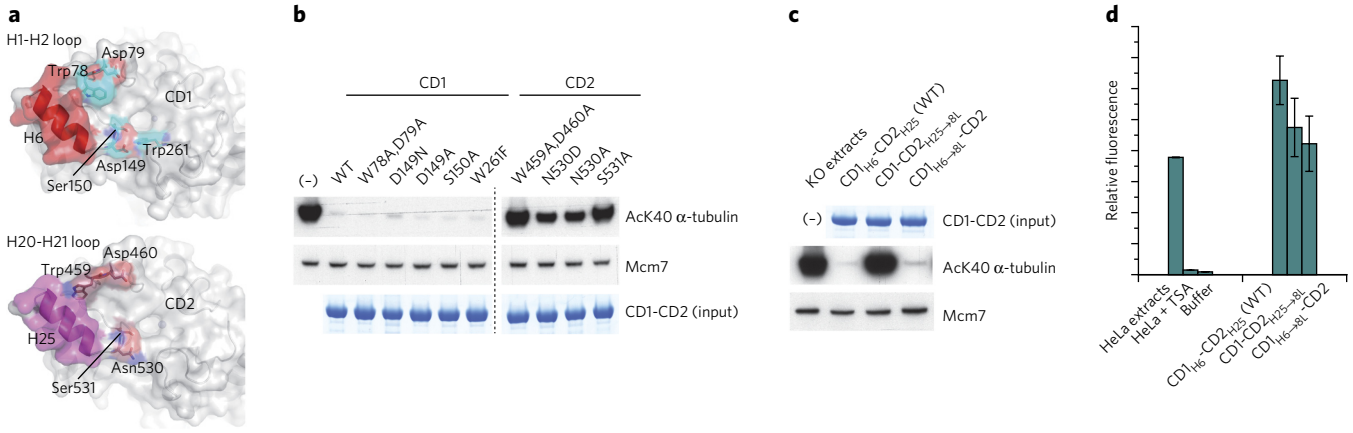


Figure 3 | Structural and molecular determinants for tubulin deacetylation. (**a**) Surface representation of CD1 and CD2 with α-helices H6 and H25, and tested amino acids indicated. (**b**) Immunoblot analysis of variant proteins in extracts from *Hdac6* knockout MEFs tested using an antibody against α-tubulin acetylated at Lys40 (AcK40), or Mcm7 to control for equal extract amount and loading. Gel at the bottom shows equal amounts of the different proteins used (full gel images are available in **Supplementary Fig. 11**). (−) denotes *Hdac6* knockout MEFs extract only (without addition of any protein). WT, wild-type CD1-CD2 sequence. (**c**) Immunoblot analysis of HDAC6 α-helix (H6 and H25) to HDAC8 loop swap changes in CD2 ($CD1_{H6\rightarrow 8L}$-CD2, $CD1\text{-}CD2_{H25\rightarrow 8L}$), were tested with HDAC6 knockout (KO) extracts as in **b** and compared to the wild-type CD1-CD2 fusion, $CD1_{H6}\text{-}CD2_{H25}$ (WT). The top gel shows equal amounts of the different proteins used (full gel images are available in **Supplementary Fig. 12**). (**d**) Enzymatic activity of purified HDAC6 proteins with the indicated changes as in **c**, tested for activity with a Fluor de Lys HDAC assay kit. HeLa extract with or without TSA addition was used as a control. Shown are mean values of three independent experiments with s.d.

Conservation of domain-domain interface

We used the above information to examine in more detail the CD1-CD2 interface (**Fig. 2a**). Several patches of conserved residues took part in the interface, and a large area was composed of less conserved residues. Important hydrophobic binding energy was provided by CD1 residues Pro306, Pro383, Leu305 and Ile405, and on the CD2 side Pro688, Ile783 and the Arg691 guanidinium group contributed the most. A key interaction was formed by the side chain stacking of His345 and His727, located at the same structural position in CD1 and CD2, with the pseudo-twofold axis running through the imidazole stacking pair. Neither the linker connecting CD1 and CD2 (418–442), nor loops H17-H18 in CD1 and H36-H37 in CD2, all of which contribute to the interface, are highly conserved (**Supplementary Fig. 2**). Nevertheless, a closer look at the linker and its interaction with CD1 and CD2 revealed that it interacted with conserved residues on both domains (for example, Pro383 on CD1; Arg605, Glu692 and Trp794 on CD2) and provided a seal between them (**Fig. 2b**).

Structural features critical for tubulin deacetylation

Because HDAC6 is a major tubulin deacetylase, it is important to understand which features endow it with the capacity to deacetylate this substrate. Calculation of the electrostatic surface potentials around each catalytic pocket highlights the mixed hydrophobic and polar character and the differences between CD1 and CD2 (**Fig. 2c**). This observation suggests the potential for differential substrate recognition by these two domains (**Fig. 3**). We hypothesized that the H1-H2 and H20-H21 loops that contain Trp78 and Asp79 in CD1, or Trp459 and Asp460 in CD2, are flexible and contribute to substrate recognition. In addition, we focused on several highly conserved amino acids, which are also around, but not in, the catalytic pocket, in particular Ser150 in CD1 and Ser531 in CD2, which correspond to Asp101 in the H6-H7 loop of HDAC8. This residue is critical for substrate or inhibitor binding by HDAC8 (ref. 42). We analyzed activities of alanine substitutions (loss of function) or conserved amino acid substitutions (gain of function) variants *in vitro* (**Fig. 3a**,**b**). To determine the activity of the mutants against the physiologic substrate α-tubulin, we incubated purified HDAC6 proteins with extracts from *Hdac6* knockout MEFs, in which α-tubulin is fully acetylated[23], and measured the resulting level of α-tubulin K40 acetylation by immunoblotting. We also assayed the enzymatic potential of the variants on the small substrate Fluor de Lys. The different point mutations in the sequence encoding CD1 did not result in compromised deacetylation activity on α-tubulin. In contrast, several point mutants in sequences encoding CD2 (the W459A,D460A double-substitution variant, or variants with single substitutions N530A, N530D or S531A) were all strongly impaired for α-tubulin deacetylation (**Fig. 3b** and **Supplementary Figs. 11** and **12**). However, when we tested these variants on the Fluor de Lys substrate, their activity was almost intact (**Supplementary Fig. 13**), suggesting that these residues are involved in substrate recognition, but not in the catalytic process.

Importance of H6 and H25 and of tandem catalytic domains

We considered that the unique H6 and H25 α-helices (**Supplementary Fig. 6**) might be critical for substrate specificity. Hence, we substituted them in CD1 or CD2 with the H6-H7 loop from HDAC8 (**Supplementary Fig. 14**). Replacement of CD2 H25 by loop H6-H7 of HDAC8 ($CD1\text{-}CD2_{H25\rightarrow 8L}$) dramatically impaired the activity on α-tubulin (**Fig. 3c**), whereas substitution of CD1 H6 had almost no detrimental effect ($CD1_{H6\rightarrow 8L}$-CD2). These proteins were similarly active when tested on Fluor de Lys substrate (**Fig. 3d**), indicating that the catalytic potential of HDAC6 was not impaired, but rather the capacity to use tubulin as a substrate was impaired. This was further demonstrated by *in vivo* experiments in which we stably reintroduced the same variant zebrafish proteins into *Hdac6* knockout MEFs and monitored tubulin acetylation by immunoblotting. The results of these experiments were identical to the *in vitro* results, and confirmed the critical role of CD2 H25 (**Supplementary Fig. 15**).

We next interrogated the overall contribution of each catalytic domain for tubulin deacetylation, using the same assay as above. Inactivating CD1 ($CD1_{H193A}$-CD2) had no impact, whereas inactivating CD2 ($CD1\text{-}CD2_{H574A}$) abolished tubulin deacetylation (**Supplementary Fig. 16a**). Furthermore, isolated CD2 deacetylated α-tubulin, but isolated CD1 did not; however, the isolated CD2 was about tenfold less active than CD1-CD2, as for human[31]. Moreover, adding increasing amounts of CD1 to a reaction containing a

fixed amount of CD2 did not influence tubulin deacetylation (**Supplementary Fig. 16b**), suggesting that under these conditions CD1 cannot enhance, or inhibit, CD2 activity. When tested on a Fluor de Lys substrate, isolated CD1 showed weak but measurable activity and also enhanced the activity of CD2 when the two domains were linked (**Supplementary Fig. 17a**,**b**). The activity of zebrafish CD1 largely depends on F202, which in other species is a tyrosine (**Supplementary Fig. 2**); conversion of F202 into a tyrosine ($CD1_{F202Y}$) strongly impaired activity (**Supplementary Fig. 17c**). The fact that zebrafish CD1 had no activity on α-tubulin but was catalytically active on Fluor de Lys suggested that it may recognize other substrates.

Stochastic deacetylation of microtubules by HDAC6

We next examined how HDAC6 deacetylates tubulin or MTs (for their preparation, see Online Methods and **Supplementary Fig. 18a–c**). We first tested whether HDAC6 can act on MTs from the ends, or whether like α-TAT it diffuses in MTs and deacetylates them stochastically. We reacted MTs with the HDAC6 tandem catalytic domains using varying enzyme amounts or for various durations; then we fixed them and analyzed by fluorescence microscopy. Addition of the HDAC6 catalytic domains led to deacetylation of the MTs over their entire length, in a dose-dependent and stochastic manner (**Fig. 4a**,**b**). We observed no preferential reaction toward the ends of MTs, even at the shortest incubation duration (**Fig. 4c**). Thus, *in vitro,* the α-tubulin Lys40 in the lumen of MTs was accessible for deacetylation by HDAC6. Next, we compared the capacity of HDAC6 to deacetylate tubulin dimers or taxol-stabilized MTs; for this, we incubated HDAC6 tandem catalytic domains with radiolabeled acetylated tubulin or MTs and used the TCA-precipitable radioactivity to determine HDAC6 activity. Deacetylation was ~2.5× more effective on tubulin heterodimers than on MTs (**Fig. 4d**), and we obtained the same results with a different experimental setup (**Supplementary Fig. 19**).

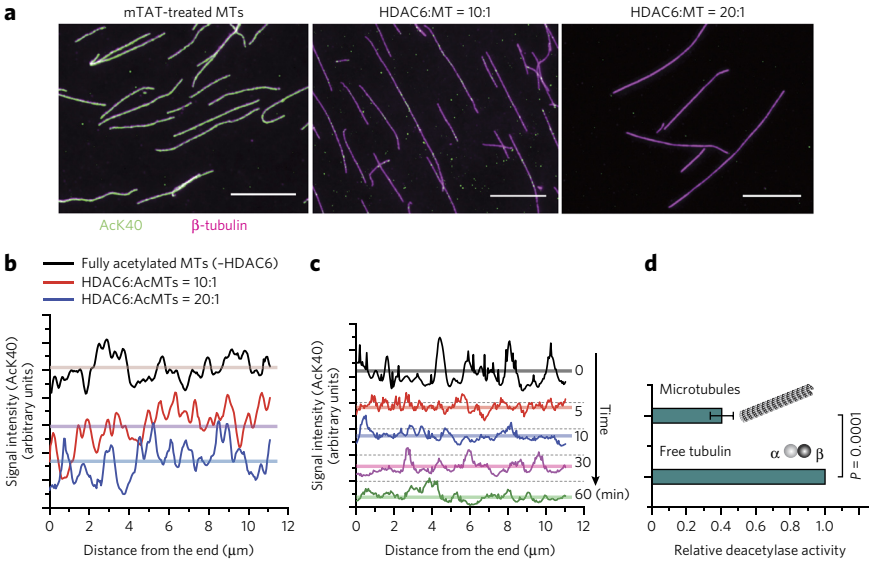


Figure 4 | HDAC6 prefered tubulin dimers, but deacetylated MTs stochastically. (**a**) Microscopy analysis of fully acetylated MTs (mTAT-treated; left) and with HDAC6 added at indicated ratios to monitor deacetylation (middle and right). Scale bars, 10 μm. Magenta staining identifies β-tubulin (MTs), and green staining corresponds to AcK40 of α-tubulin. (**b**) Quantification of fluorescence relative to that of fully acetylated MTs in **a** for an average of 20 line scans of microtubule ends; all scanned MTs were longer than 10 μm and initial MTs had been fully acetylated by α-TAT. Fluorescence intensity of fully acetylated MTs is shown by the black line. Fluorescence intensity of MTs deacetylated by HDAC6 treatment is shown by red (HDAC6:AcMTs = 10:1) and blue (HDAC6:AcMTs = 20:1) lines, respectively. Horizontal lines show respective average fluorescence intensity. (**c**) Time course deacetylation experiment on MTs, with scans indicating AcK40 signal intensity along the MTs at indicated time points after HDAC6 addition. Dashed lines indicate *x* axis for each time point. Line scans from each time point are staggered vertically for clarity. (**d**) Deacetylation activity on MTs vs. free tubulin dimers. Shown are mean values of three independent experiments with s.d.; *P* value is based on a two-tailed Student's *t*-test. Deacetylation activity on free tubulin dimers was set to 1.

Structure of CD1-CD2 in complex with NextA

Although no crystal structure of HDAC6 was available, several HDAC6 selective inhibitors have been developed in recent years[40,43,44]. To gain structural insight into HDAC6 selective inhibition, we determined the structure of the HDAC6 tandem catalytic domains in complex with NextA, an inhibitor with high selectivity for HDAC6 (ref. 40). NextA features a classical hydroxamate zinc-binding group (ZBG) with a benzylic linker connected to a urea-based cap group consisting of a second benzyl and an n-butyl moiety (**Supplementary Fig. 20a**). NextA bound to CD1 and CD2 active sites with distinct characteristics when compared to our TSA structures, giving insight into selective inhibition (**Fig. 5a** and **Supplementary Fig. 20b–d**).

Human HDAC6 homology model and NextA selectivity

Using our zebrafish HDAC6 structures we computed CD1 and CD2 homology models of the human ortholog to understand selectivity of NextA for human HDAC6. Zebrafish CD1 has two critical positions His82 and Phe202, which are Phe105 and Tyr225 in the human protein (**Supplementary Fig. 21a**). As mentioned above, mutation of zebrafish Phe202 into tyrosine resulted in a strong reduction in activity (**Supplementary Fig. 17c**). Human CD2 in contrast had only two amino acid changes located at the periphery of the pharmacophore: zebrafish residues Asn530 and Asn645 are Asp567 and Met682 in the human protein, and all residues found to interact with (*S*)-TSA or NextA are fully conserved (**Supplementary Fig. 21b**). This indicates that the structure of zebrafish CD2 and the corresponding homology model of the human protein may be valid to understand selective NextA inhibition of HDAC6 over other HDAC isoforms.

Superposition of all HDAC isoform structures with zebrafish HDAC6 CD2 and the corresponding human homology model revealed important differences between isoform-specific pharmacophores influencing inhibitor selectivity (**Fig. 5a–c** and **Supplementary Fig. 21c**). Owing to the unique position of helix H25 and the conformation of the following loop, only HDAC6 had a large open basin ~14 Å wide. NextA selectivity for HDAC6 thus seems to come from (i) the isoform-specific shape and height of the rim between the wide HDAC6 basin and the acetylated lysine binding channel where the NextA benzyl cap group docked, and (ii) the bulkiness of residues occupying the basin in other isoforms. Owing to the rigid nature of NextA and the 90° angle between the linker and the urea-benzyl cap, these isoform-specific steric constraints will determine how far the short NextA benzylic linker can reach into the cavity and whether the ZBG can complex the Zn^{2+} ion favorably. Given these structural features, only HDAC6 provided sufficient space in this region to allow tight NextA binding,

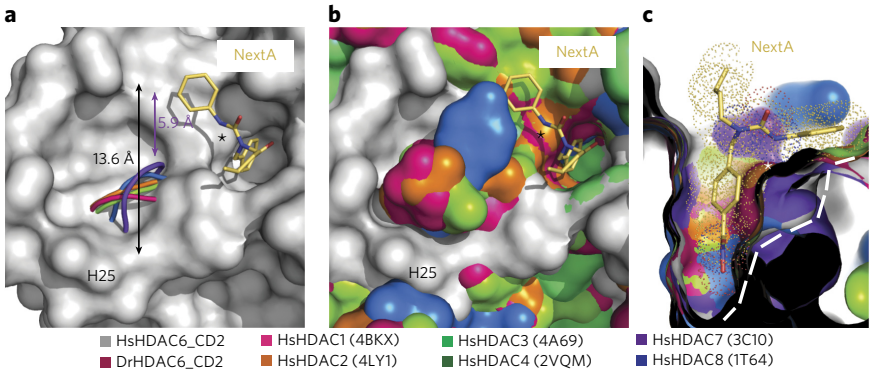


Figure 5 | HDAC6-specific inhibitor binding. (**a**–**c**) Superposition of HDAC isoforms with zebrafish (Dr) HDAC6 CD2 (dark red) and its human (Hs) homology model (gray) to highlight differences in the active site architecture. Human HDAC6 CD2 is shown in surface representation, and all other isoforms are cartoon models (**a**). NextA is shown as sticks in gold; arrows highlight the width of the basin adjacent to the acetylated lysine binding channel for HDAC6 (black) and HDAC7 (violet). HDAC6 helix H25 is labeled. The asterisk marks the rim under the NextA cap group which defines selective inhibitor access in different isoforms. In **b**, all HDAC structures are depicted in surface representation. Cross-section through the acetylated lysine binding channel under the NextA cap group (**c**). All HDAC structures are shown in surface mode. The white dashed line underlines the lowered rim of the channel for HDAC6–8, while all other isoforms have a much more elevated rim (asterisk) and therefore restricted access to the active site. Van der Waals radii for NextA atoms are shown as dots.

whereas other isoforms restricted binding either with bulky side chains occupying the basin or with an elevated rim hindering the ZBG from reaching the Zn^{2+} ion. Similarly, only HDAC6 provided the unique position of H25 and conformation of the following loop, which, together with residues located in loop H29-H30, seemed to provide superior binding energy for (*S*)-TSA, explaining selectivity over other HDAC isoforms (**Supplementary Fig. 21d**).

DISCUSSION

The crystal structure of the HDAC6 tandem catalytic domains revealed that the two domains interact over a large surface area, and that both catalytic sites point outside and are accessible to substrates. The interdomain linker, which varies in length between species, was at the outside of the complex and sealed the two catalytic domains. Previous mutagenesis studies had shown an important role of the linker for optimal activity of HDAC6 (ref. 32). Moreover the linker region is essential for interaction between human or mouse HDAC6 and dynein motor proteins[6,9]. Given the position of the linker observed in our structure, it is conceivable that it forms part of an interface interacting with dynein, possibly together with additional residues in the catalytic domains.

The structure of each domain in complex with inhibitors revealed features of the catalytic pockets. The binding of (*R*)-TSA and (*S*)-TSA to CD1 and CD2 was highly similar for the hydroxamate moiety, but crystal lattice constraints favored a packing of CD2 in complex with (*S*)-TSA where the CAP group orientation differed compared with (*R*)-TSA binding to CD1. Our observations did not indicate a preferential binding of (*S*)-TSA to CD2, but were intriguing, as early reports had suggested that this enantiomer is inactive[37,38]. We found that *in vitro* both pure (*S*)-TSA and (*R*)-TSA inhibited similarly HDAC6. In contrast, when tested on all human HDACs, the two forms showed distinct inhibitory profiles, and (*S*)-TSA appeared to be a HDAC6-selective inhibitor, with ~20-fold selectivity for that isoform. This observation may open up avenues for the generation of new HDAC6-specific inhibitors. Furthermore, in MEF cells (*S*)-TSA was active, albeit less than (*R*)-TSA, and also showed moderate selectivity. The reduced activity of (*S*)-TSA on the main class I HDACs may explain why it had initially been considered to be biologically inactive[37,38].

Mutational analysis revealed that helix H25 and the loop H20-H21 in CD2 were critical for deacetylation of Lys40 on α-tubulin, but not for the small substrate Fluor de Lys; together with Asn530 and Ser531 they form the recognition platform for the α-tubulin loop encompassing Lys40.

Early studies showed that deacetylation of MTs correlates with their depolymerization *in vivo*[45] and in *C. elegans* mutation of the acetyltransferase MEC-17 (the homolog of α-TAT) affects MTs integrity[46,47]. In cultured cells and mouse organs tubulin acetylation is usually low due to the action of HDAC6 and possibly SIRT2 (refs. 20,21,23). Ablation of HDAC6 in mice or cells leads to an almost complete acetylation of α-tubulin, indicating that the balance between HDAC6 and α-TAT is critical to maintain physiological levels of this modification. Recently molecular and structural studies described how α-TAT acetylates MTs, in preference over tubulin; thereby, the α-TAT enzyme enters the lumen of MTs and acetylates them stochastically[19]. When HDAC6 is tested in similar assays it can also deacetylate Lys40 in the lumen of MTs, but the preferred substrate is unpolymerized tubulin, as also observed in ref. 26.

Although zebrafish HDAC6 showed some differences compared to the human enzyme (for example, the activity of CD1 on Fluor de Lys substrate) our overall analysis indicated that it is a valid model to describe the human enzyme. Our structure of CD1-CD2 in complex with the HDAC6-specific inhibitor NextA and homology modeling of the human CD2 domain help to understand selective inhibition of HDAC6. The unique position of H25 and the conformation of the following loop provided selectivity for NextA, (*S*)-TSA and likely also other HDAC6-specific inhibitors. The results presented here will be useful to better understand the biology of HDAC6 and to accelerate drug development.

Received 4 December 2015; accepted 22 June 2016;
published online 25 July 2016

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 5G0G (CD1-TSA), 5G0H (CD2-TSA), 5G0I (CD1-CD2 NextA, cleaved linker), 5G0J (CD1-CD2 NextA, linker intact) and 5G0F (ZnF-UBP).

References

1. Yang, X.J. & Seto, E. The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. *Nat. Rev. Mol. Cell Biol.* **9**, 206–218 (2008).
2. Hubbert, C. *et al.* HDAC6 is a microtubule-associated deacetylase. *Nature* **417**, 455–458 (2002).
3. Kovacs, J.J. *et al.* HDAC6 regulates Hsp90 acetylation and chaperone-dependent activation of glucocorticoid receptor. *Mol. Cell* **18**, 601–607 (2005).
4. Zhang, X. *et al.* HDAC6 modulates cell motility by altering the acetylation level of cortactin. *Mol. Cell* **27**, 197–213 (2007).
5. Boyault, C. *et al.* HDAC6 controls major cell response pathways to cytotoxic accumulation of protein aggregates. *Genes Dev.* **21**, 2172–2181 (2007).

6. Kawaguchi, Y. *et al.* The deacetylase HDAC6 regulates aggresome formation and cell viability in response to misfolded protein stress. *Cell* **115**, 727–738 (2003).
7. Kwon, S., Zhang, Y. & Matthias, P. The deacetylase HDAC6 is a novel critical component of stress granules involved in the stress response. *Genes Dev.* **21**, 3381–3394 (2007).
8. de Zoeten, E.F. *et al.* Histone deacetylase 6 and heat shock protein 90 control the functions of Foxp3(+) T-regulatory cells. *Mol. Cell. Biol.* **31**, 2066–2078 (2011).
9. Banerjee, I. *et al.* Influenza A virus uses the aggresome processing machinery for host cell entry. *Science* **346**, 473–477 (2014).
10. Aldana-Masangkay, G.I. & Sakamoto, K.M. The role of HDAC6 in cancer. *J. Biomed. Biotechnol.* **2011**, 875824 (2011).
11. Seidel, C., Schnekenburger, M., Dicato, M. & Diederich, M. Histone deacetylase 6 in health and disease. *Epigenomics* **7**, 103–118 (2015).
12. Lawson, J.L.D. & Carazo Salas, R.E. Microtubules: greater than the sum of the parts. *Biochem. Soc. Trans.* **41**, 1736–1744 (2013).
13. Janke, C. & Bulinski, J.C. Post-translational regulation of the microtubule cytoskeleton: mechanisms and functions. *Nat. Rev. Mol. Cell Biol.* **12**, 773–786 (2011).
14. Yu, I., Garnham, C.P. & Roll-Mecak, A. Writing and Reading the Tubulin Code. *J. Biol. Chem.* **290**, 17163–17172 (2015).
15. Nogales, E., Whittaker, M., Milligan, R.A. & Downing, K.H. High-resolution model of the microtubule. *Cell* **96**, 79–88 (1999).
16. Soppina, V., Herbstman, J.F., Skiniotis, G. & Verhey, K.J. Luminal localization of α-tubulin K40 acetylation by cryo-EM analysis of fab-labeled microtubules. *PLoS One* **7**, e48204 (2012).
17. Kormendi, V., Szyk, A., Piszczek, G. & Roll-Mecak, A. Crystal structures of tubulin acetyltransferase reveal a conserved catalytic core and the plasticity of the essential N terminus. *J. Biol. Chem.* **287**, 41569–41575 (2012).
18. Shida, T., Cueva, J.G., Xu, Z., Goodman, M.B. & Nachury, M.V. The major alpha-tubulin K40 acetyltransferase alphaTAT1 promotes rapid ciliogenesis and efficient mechanosensation. *Proc. Natl. Acad. Sci. USA* **107**, 21517–21522 (2010).
19. Szyk, A. *et al.* Molecular basis for age-dependent microtubule acetylation by tubulin acetyltransferase. *Cell* **157**, 1405–1415 (2014).
20. Zhang, Y. *et al.* HDAC-6 interacts with and deacetylates tubulin and microtubules in vivo. *EMBO J.* **22**, 1168–1179 (2003).
21. North, B.J., Marshall, B.L., Borra, M.T., Denu, J.M. & Verdin, E. The human Sir2 ortholog, SIRT2, is an NAD+-dependent tubulin deacetylase. *Mol. Cell* **11**, 437–444 (2003).
22. Nahhas, F., Dryden, S.C., Abrams, J. & Tainsky, M.A. Mutations in SIRT2 deacetylase which regulate enzymatic activity but not its interaction with HDAC6 and tubulin. *Mol. Cell. Biochem.* **303**, 221–230 (2007).
23. Zhang, Y. *et al.* Mice lacking histone deacetylase 6 have hyperacetylated tubulin but are viable and develop normally. *Mol. Cell. Biol.* **28**, 1688–1701 (2008).
24. Zilberman, Y. *et al.* Regulation of microtubule dynamics by inhibition of the tubulin deacetylase HDAC6. *J. Cell Sci.* **122**, 3531–3541 (2009).
25. Matsuyama, A. *et al.* In vivo destabilization of dynamic microtubules by HDAC6-mediated deacetylation. *EMBO J.* **21**, 6820–6831 (2002).
26. Zhao, Z., Xu, H. & Gong, W. Histone deacetylase 6 (HDAC6) is an independent deacetylase for alpha-tubulin. *Protein Pept. Lett.* **17**, 555–558 (2010).
27. Ageta-Ishihara, N. *et al.* Septins promote dendrite and axon development by negatively regulating microtubule stability via HDAC6-mediated deacetylation. *Nat. Commun.* **4**, 2532 (2013).
28. Falkenberg, K.J. & Johnstone, R.W. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat. Rev. Drug Discov.* **13**, 673–691 (2014).
29. Schölz, C. *et al.* Acetylation site specificities of lysine deacetylase inhibitors in human cells. *Nat. Biotechnol.* **33**, 415–423 (2015).
30. Santo, L. *et al.* Preclinical activity, pharmacodynamic, and pharmacokinetic properties of a selective HDAC6 inhibitor, ACY-1215, in combination with bortezomib in multiple myeloma. *Blood* **119**, 2579–2589 (2012).
31. Zou, H., Wu, Y., Navre, M. & Sang, B.C. Characterization of the two catalytic domains in histone deacetylase 6. *Biochem. Biophys. Res. Commun.* **341**, 45–50 (2006).
32. Zhang, Y., Gilquin, B., Khochbin, S. & Matthias, P. Two catalytic domains are required for protein deacetylation. *J. Biol. Chem.* **281**, 2401–2404 (2006).
33. Ouyang, H. *et al.* Protein aggregates are recruited to aggresome by histone deacetylase 6 via unanchored ubiquitin C termini. *J. Biol. Chem.* **287**, 2317–2327 (2012).
34. Finnin, M.S. *et al.* Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. *Nature* **401**, 188–193 (1999).
35. Lombardi, P.M., Cole, K.E., Dowling, D.P. & Christianson, D.W. Structure, mechanism, and inhibition of histone deacetylases and related metalloenzymes. *Curr. Opin. Struct. Biol.* **21**, 735–743 (2011).
36. Schuetz, A. *et al.* Human HDAC7 harbors a class IIa histone deacetylase-specific zinc binding motif and cryptic deacetylase activity. *J. Biol. Chem.* **283**, 11355–11363 (2008).
37. Yoshida, M., Hoshikawa, Y., Koseki, K., Mori, K. & Beppu, T. Structural specificity for biological activity of trichostatin A, a specific inhibitor of mammalian cell cycle with potent differentiation-inducing activity in Friend leukemia cells. *J. Antibiot. (Tokyo)* **43**, 1101–1106 (1990).
38. Yoshida, M., Kijima, M., Akita, M. & Beppu, T. Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A. *J. Biol. Chem.* **265**, 17174–17179 (1990).
39. Cosner, C.C. *et al.* Evolution of Concise and Flexible Synthetic Strategies for Trichostatic Acid and the Potent Histone Deacetylase Inhibitor Trichostatin A. *Eur. J. Org. Chem.* 162–172 (2013).
40. Bergman, J.A. *et al.* Selective histone deacetylase 6 inhibitors bearing substituted urea linkers inhibit melanoma cell growth. *J. Med. Chem.* **55**, 9891–9899 (2012).
41. Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, W529–W533 (2010).
42. Dowling, D.P., Gantt, S.L., Gattis, S.G., Fierke, C.A. & Christianson, D.W. Structural studies of human histone deacetylase 8 and its site-specific variants complexed with substrate and inhibitors. *Biochemistry* **47**, 13554–13563 (2008).
43. Butler, K.V. *et al.* Rational design and simple chemistry yield a superior, neuroprotective HDAC6 inhibitor, tubastatin A. *J. Am. Chem. Soc.* **132**, 10842–10846 (2010).
44. Silvestri, L., Ballante, F., Mai, A., Marshall, G.R. & Ragno, R. Histone deacetylase inhibitors: structure-based modeling and isoform-selectivity prediction. *J. Chem. Inf. Model.* **52**, 2215–2235 (2012).
45. Black, M.M., Baas, P.W. & Humphries, S. Dynamics of alpha-tubulin deacetylation in intact neurons. *J. Neurosci.* **9**, 358–368 (1989).
46. Cueva, J.G., Hsin, J., Huang, K.C. & Goodman, M.B. Posttranslational acetylation of α-tubulin constrains protofilament number in native microtubules. *Curr. Biol.* **22**, 1066–1074 (2012).
47. Topalidou, I. *et al.* Genetically separable functions of the MEC-17 tubulin acetyltransferase affect microtubule organization. *Curr. Biol.* **22**, 1057–1065 (2012).

Acknowledgments

We thank J. Spector and A. Roll-Mecak for sharing information about total internal reflection fluorescence microscopy experiments and discussions, S. Weiler and U. Schopf and U. Rass for helpful discussions, K. Verhey for a technical suggestion, L. Gelman for help with microscopy analysis, and D. Klein and J. Seebacher for mass spectrometry analysis to identify domain boundaries. Part of this work was performed at beamlines X10SA and X06DA of the Swiss Light Source. We thank G. Matthias, C. Cao and M. Regenass for helpful technical assistance, and O. Truee, R.G. Clerc and all the members of the Matthias laboratory for fruitful discussions. This work was supported by the Novartis Research Foundation and M.S. was also partly supported by a Fellowship from the Nakajima Foundation. Work performed at the University of Notre Dame is supported by the Ara Parseghian Medical Research Foundation, the National Institutes of Health (1R01NS092653), the Warren Family Center for Drug Discovery and Development and the Department of Chemistry and Biochemisry. X.W., B.J.M. and P.H. thank J. Zajicek and J. Pontius for NMR microscopy support and discussion.

Author contributions

Y.M., H.G. and P.M. designed experiments; Y.M. performed biochemical and microscopy experiments, Y.M., J.J.K., L.W. and M.S. prepared and purified proteins and performed assays; M.S. and L.W. performed cellular inhibitor assays; J.J.K. crystallized proteins; X.W., B.J.M. and P.H. synthesized and purified (*R*)- and (*S*)-TSA; H.G. and J.J.K. collected diffraction data and H.G. determined crystal structures; D.H. analyzed mass spectrometry data; Y.M., H.G. and P.M. wrote the manuscript; P.M. oversaw the work. All authors contributed to the final manuscript.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Any supplementary information, chemical compound information and source data are available in the online version of the paper. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to P.M.

ONLINE METHODS

Cloning of HDAC6 constructs. A full-length *Danio rerio* HDAC6 cDNA clone was made by synthesizing the first 495 base pairs (bp) of the coding region (GeneArt) and fusing it to the partial cDNA IMAGE clone 7051100 (Source BioScience) via In-Fusion cloning (ClonTech). The translated sequence corresponds to Uniprot F8W4B7. PCR products were cloned into pOPINF, which introduces sequence encoding an N-terminal 6×His tag and a 3C protease cleavage site before the sequence encoding HDAC6; in some cases, pOPINM vectors containing sequence encoding an N-terminal 6×His tag and a maltose binding protein (MBP) tag before the 3C protease cleavage site were used.

Expression and purification of HDAC6 proteins. HDAC6 CD1-CD2 (encompassing amino acid residues 25–831, or 40–831) and HDAC6 CD1 (40–418) were expressed in Sf9 insect cells using the FlashBAC baculovirus system. HDAC6 CD1-CD2 protein was extracted from a baculovirus-infected Sf9 cell pellet by thoroughly resuspending the cells in ice-cold nickel lysis buffer (50 mM Tris, pH 7.5, 200 mM NaCl, 20 mM imidazole, 5% glycerol, 2 mM TCEP, 0.2% Tween-20), freshly supplemented with Complete EDTA-free protease inhibitors (Roche) and Benzonase (Sigma). After 20 min on ice the lysate was centrifuged at 30,000*g* for 30 min at 4 °C. The clarified soluble lysate was incubated in batch mode with Ni-NTA IMAC agarose (Qiagen), and then transferred into a 10 ml Econo-Pac column (Bio-Rad) for washing with nickel wash buffer (50 mM Tris, pH 7.5, 200 mM NaCl, 20 mM imidazole, 5% glycerol, 2 mM TCEP). The target protein was eluted in nickel wash buffer containing 125 mM imidazole. The imidazole concentration in the eluate was adjusted to 20 mM. After an overnight digestion at 4 °C with His-tagged 3C protease, the cleaved HDAC6 was further purified over Ni-NTA agarose resin. The flow-through and wash fractions containing untagged HDAC6 were pooled, concentrated and separated using an AKTA Pure system (GE healthcare) with a HiLoad 16/600 Superdex 200 (GE Healthcare) gel filtration column equilibrated in 20 mM Tris, pH 7.5, 200 mM NaCl, 2 mM TCEP, 0.02% NaN_3. In some cases, a Sephacryl S-300 16/60 gelfiltration column (GE Healthcare) run on a DUO FLOW system (Bio-Rad) was used. Protein fractions were analyzed on a 4–12% Bis-Tris NuPAGE gel and pure fractions containing HDAC6 CD1-CD2 were pooled and concentrated to 11 mg/ml. Gels were stained with InstantBlue. HDAC6 CD1 protein was purified as described above for the HDAC6 CD1-CD2 protein, with the following modifications. Two gel-filtration steps were performed on S200 in 20 mM Tris, pH 7.5, 200 mM NaCl, 2 mM TCEP, 0.02% NaN_3. The first gel filtration run was after the initial Ni-NTA affinity step and the second gel filtration run was as a final polishing step. HDAC6 CD1 protein was concentrated to 10 mg/ml.

HDAC6 ZnF-UBP (974–1081) was expressed in *E. coli* BL21 (DE3) cells and induced with 0.5 mM IPTG at 20 °C for 20 h. *E. coli* BL21 (DE3) cells expressing HDAC6 ZnF-UBP were pelleted, resuspended in lysis buffer (50 mM Tris, pH7.5, 500 mM NaCl, 20 mM imidazole, 1 mM TCEP, 0.2% Tween-20) then rapidly frozen on dry-ice and stored at −80 °C. The frozen cell suspension was thawed at room temperature and supplemented with Complete EDTA-free protease inhibitors (Roche) and 3 U/ml Benzonase (Sigma), before passing through an Avestin EmulsiFlex-C3 cell disruptor. The clarified soluble lysate was incubated with Ni-NTA superflow resin (Qiagen) in batch mode and the bound protein was eluted in 50 mM Tris, pH 7.5, 500 mM NaCl, 500 mM imidazole. The protein was digested overnight with His-tagged 3C protease while dialyzing against 20 mM Tris, pH 7.5, 200 mM NaCl, 20 mM imidazole, 1 mM TCEP, 0.02% NaN_3 at 4 °C. The dialyzed protein was filtered through a 0.22 µm filter and then purified over Ni-NTA superflow resin. Untagged HDAC6 ZnF-UBP protein was collected in the flow-through fraction and concentrated before separating on a Superdex 75 HiLoad 16/60 (GE Healthcare) gel filtration column equilibrated in 20 mM Tris, pH 7.5, 200 mM NaCl, 2 mM TCEP and 0.02% NaN_3. Fractions corresponding to the pure HDAC6 ZnF-UBP protein were pooled and concentrated to 14 mg/ml.

Crystallization of HDAC6 proteins. Nanoliter crystallization experiments were performed with a Phoenix dispensing robot (Art Robbins) using the sitting-drop vapor diffusion method at 20 °C. Drops with HDAC6 CD1-CD2 (40-831) protein at 14 mg/ml and 1.3 mM Nexturastat A (NextA, BioVision), an HDAC6-specific inhibitor, crystallized after 5 d in 3.3 M sodium formate. The single crystals were harvested after 15 d and cryoprotected in 3.3 M sodium formate, 0.1 M Tris, pH 7.5, 17% glycerol, 2 mM NextA and 4% DMSO. These crystals contained CD1 joined to CD2 via its internal intact linker. Crystallization experiments using HDAC6 CD1-CD2 (25–831) protein at 11 mg/ml and 0.66 mM NextA, yielded thick plate crystals after 40 days in 23.2% PEG3350, 0.1 M KCl. These crystals contained CD1 associated with CD2 although the linker region was proteolytically cleaved. The crystals were cryoprotected in mother liquid containing 17% ethylene glycol, 0.6 mM NextA and 1.3% DMSO. HDAC6 CD1 at 10 mg/ml was incubated with 1 mM TSA (*R* form, Sigma), a pan-HDAC inhibitor. Crystals did not appear in the absence of HDAC inhibitor. Long thick plate crystals appeared after 2 d in 23.2% PEG 3350, 0.2 M $MgCl_2$, 0.1 M HEPES, pH 7.0 and Silver Bullet additive A7 (Hampton Research). The crystals were harvested after 3 weeks and cryoprotected in mother liquor containing 20% ethylene glycol, 4 mM TSA (*R* form, Sigma) and 4% DMSO. Crystals containing HDAC6 CD2 were formed from crystallization experiments set up with HDAC6 CD1-CD2 at 9 mg/ml with 0.5 mM TSA (racemic mixture, MBL) in 15% PEG 3350, 0.1 M KCl. Long thick plate crystals appeared after 13 days and were harvested two days later and cryoprotected in mother liquor containing 25% ethylene glycol and 0.5 mM TSA (racemic mixture, MBL) and 1.7% DMSO. Two crystals used for data collection were dissolved in H_2O and analyzed by SDS-PAGE and mass-spectrometry. Crystals contained only the CD2 domain likely resulting from slow proteolysis of the CD1-CD2 protein during crystallization. Crystallization experiments with HDAC6 ZnF-UBP at 14.4 mg/ml yielded many poor quality crystals soon after the trays were dispensed. By slowing down the nucleation process, the best crystals appeared after several months in 1 M Li_2SO_4, 5 mM $NiCl_2$, 0.1 M Tris, pH 8.5. Crystals were harvested and cryoprotected in 2 M Li_2SO_4.

Size-exclusion chromatography with multi-angle light scattering. Purified zebrafish HDAC6 proteins were concentrated to 1–5 mg/ml and filtered through a 0.1 µM Amicon filter before injection. In all, 38 µl of each protein was separated on a Superdex 200 10/300 GL gel-filtration column (GE Healthcare) equilibrated in 20 mM Tris, pH 7.5, 200 mM NaCl, 1 mM TCEP, 0.02% NaN_3 at a flow rate of 0.65 ml/min. Light scattering was recorded on an in-line miniDAWN TREOS three-angle light scattering detector (Wyatt Technology) and protein concentration detected with an in-line Optilab Trex refractive index detector. The weight-averaged molecular mass of material contained in chromatographic peaks was determined using ASTRA 6 software (Wyatt Technology).

Data collection and structure solution. X-ray data collection was carried out at the SLS PX-II/III beamlines in Villigen, Switzerland. CD1-TSA and CD2-TSA crystals belonged to space group $C222_1$ and $P2_1$, respectively, (one chain per a.u. in both cases) and diffracted to 1.50 (λ = 0.978 Å) and 1.60 Å (λ = 1.000 Å). CD1-CD2 (cleaved linker) in complex with NextA crystallized in space group C2 (two chains per a.u.) and crystals diffracted to 2.00 Å (λ = 1.000). CD1-CD2 crystals with intact interdomain linker and NextA diffracted to 2.88 Å (λ = 1.000) and belonged to space group $P3_221$ with one molecule per a.u. and a solvent content of ~79%. The ZnF-UBP domain crystallized in space group I23 with one molecule per a.u. and crystals diffracted to 1.9 Å resolution (λ = 1.000). Diffraction data for all projects was integrated and scaled using the XDS program package[48], except for the anisotropic $P3_221$ CD1-CD2 (NextA, intact linker) diffraction data which was processed with AutoPROC[49].

CD1-TSA and CD2-TSA structures were solved by the molecular replacement method with PHASER[50] using homology models of respective zebrafish (Dr)HDAC6 domains. Phases from molecular replacement solutions were calculated and used for automatic model building with BUCCANEER[51]. Structures were then manually completed and further improved by the crystallographic simulated annealing routine followed by individual B-factor refinement in PHENIX[52]. The CD1-TSA structure was finalized by alternating rounds of rebuilding in COOT[53] and refinement in PHENIX using individual anisotropic B-factor refinement as this lowered R_{free} by more than 1% compared to isotropic ADP treatment.

The CD2-TSA structure was finalized by several rounds of manual rebuilding in COOT and refinement in BUSTER[54] using TLS and individual isotropic B-factor methods. Structures of CD1-CD2 (cleaved linker) and CD1-CD2

(intact linker) in complex with NextA were solved by molecular replacement using high resolution CD1 and CD2 structures as search models. Both CD1-CD2 structures were finalized by alternating rounds of rebuilding in COOT and refinement in BUSTER with and without TLS refinement, respectively. Map sharpening implemented in COOT was used to enhance details for modeling the CD1-CD2 (intact linker) structure. The ZnF-UBP structure was solved by molecular replacement using PDB entry 3C5K as search model and the structure was completed by iterating rounds of manual rebuilding in COOT and refinement in BUSTER.

Metal ions were modeled considering crystallization conditions and peak heights in anomalous difference Fourier electron density maps. The metal ion bound to the CD1 active site in the CD1-CD2 NextA (intact linker) structure did not display any significant peak in the anomalous difference Fourier electron density map and metal ligand distances refined to values > 2.7 Å with notable rearrangement of metal position and ligand orientations owing to a crystal contact with the Lys57 side chain from a symmetry related molecule which is bound in the active site channel. Therefore, Zn^{2+} binding was very unlikely and a K^+ ion was modeled at this position instead which refined well at 100% occupancy (no $mF_o - DF_c$ electron density peak at ± 3 σ) matching B-factor values of the environment. Contrary, the active site in CD2 of this structure (complexed by NextA) displayed a large peak in the anomalous difference Fourier electron density map (14.9 σ) with much shorter metal ligand distances confirming binding of a Zn^{2+} ion.

Metal sites were validated using the CheckMyMetal server (http://csgid.org/csgid/metal_sites) and ligand restraints were generated with the Grade web server (http://grade.globalphasing.org). Final structures were validated using COOT. Ramachandran-plot statistics: CD1-TSA: allowed 99.3%, outliers 0.7%; CD2-TSA: allowed 99.4%, outliers 0.6%, CD1-CD2 NextA (cleaved linker): allowed 99.7%, outliers 0.3%; CD1-CD2 NextA (intact linker): allowed 99.1%, outliers 0.9%; ZnF-UBP: allowed 99.0%, outliers 1.0%. Structural images for figures were prepared with PyMOL[55].

Homology modeling of human HDAC6 CD1 and CD2 domains. HHPRED[56] was used to generate a large multiple sequence alignment of HDAC6 orthologous protein sequences. Aligned zebrafish and human HDAC6 sequences from this alignment were used as input to the modeler software[57] together with high-resolution HDAC6 CD1 or CD2 structures from zebrafish as templates. TSA inhibitors and Zn^{2+} ions were included in the modeling calculations, and treated as rigid bodies. 100 models for each domain were generated and the best model was chosen according to lowest values for the modeler objective function and quality of the Ramachandran plot.

Deacetylase assays with HDAC6 knockout MEFs extracts. Microtubule deacetylation activity was measured with *Hdac6* knockout cell[23] extracts followed by immunoblotting analysis. Extracts were prepared by lysing cells on ice for 30 min with CSK buffer (10 mM PIPES, pH6.8, 300 mM sucrose, 100 mM NaCl, 3 mM $MgCl_2$, 1 mM EGTA, 0.1% Triton X-100) containing Complete protease inhibitor cocktail (Roche). The soluble supernatant containing most of tubulin was collected by centrifugation for 5 min at 13,200 r.p.m. at 4 °C and protein concentration was measured by Bradford assay.

30 µg knockout extracts were mixed with different amounts of purified HDAC6 catalytic domains, and incubated for 1 h at 37 °C. The reaction was stopped by adding SDS sample buffer and proteins were loaded onto 4–12% NuPAGE gels (Invitrogen), transferred onto polyvinylidene fluoride (PVDF) membranes (Millipore) and detected with specific antibodies (anti-AcK40 (BML-SA452-0100, Enzo, 1:3,000) and Mcm7 (47DC141(Ab2360), Abcam, 1:3,000).

HDAC activity assays and IC_{50} determinations of (*R*)-TSA and (*S*)-TSA on HDAC1–11. Enzymatic characterization (Km determinations, **Supplementary Fig. 3C**) of zebrafish HDAC6 proteins (CD1-CD2; $CD1_{H193A}$-CD2; CD1-$CD2_{H574A}$; $CD1_{H193A}$- $CD2_{H574A}$) was done with 50 nM of purified protein and increasing amounts of Fluor de Lys substrate, using an HDAC assay kit from Enzo biochemicals and following the manufacturer's instructions. Human HDAC6 protein (HsHDAC6 FL) was obtained from Reaction Biology Corporation. The fluorescence intensity was detected with a Spectromax Gemini plate reader (Molecular Devices). Curve fitting was done using GraphPad (GraphPad Software). The values expressed are the average of duplicate independent trials ± s.d.

IC_{50} of (*R*)- and (*S*)-TSA on human HDACs as well as on zebrafish HDAC6 were determined by Reaction Biology Corporation. Fluorogenic peptide from p53 residues 379-382 (RHKK(Ac)AMC) was used for HDAC1, 2, 3, 6 and zebrafish HDAC6. Fluorogenic HDAC class IIa substrate (trifluoroacetyl Lysine) was used for HDAC4, 5, 7, 9, and 11. Fluorogenic peptide from p53 residues 379–382 (RHK(Ac)K(Ac)AMC) was used for HDAC8. The assay buffer contained Tris-HCL, pH 8.0, 127 mM NaCl, 2.7 mM KCl, 1 mM $MgCl_2$, 1 mg/mL BSA and 1% DMSO. Inhibitors were diluted in DMSO, preincubated with the enzyme for 10 min, after which substrate was added and the reaction allowed to proceed for 2 h at 30 °C. The reaction was terminated by addition of TSA and developer. Dose-response curves were generated by serial threefold dilution of compound to generate 10-dose plots; curve-fitting was done with GraphPad. IC_{50} values were derived from the plots and the values are expressed as the average of duplicate determinations ± s.d.

Cellular assays with (*R*)-TSA and (*S*)-TSA. MEFs were treated with the different inhibitors ((*R*)-TSA, (*S*)-TSA, TSA (Sigma) and NextA) for 12 h. The cells were washed with PBS and lysed in RIPA buffer (50 mM Tris–HCl, pH8.0, 500 mM sodium chloride, 1 mM EDTA, 1.0% NP-40, 0.5% sodium deoxycholate, 0.1% SDS and Complete EDTA-free protease inhibitors (Roche)). After removal of the insoluble fraction by centrifugation, 20-30 µg protein extract was boiled for 10 min in LDS sample buffer (Invitrogen), and separated on 4–12% NuPAGE gels (Invitrogen). Proteins were transferred onto PVDF membranes (Immobilon-P, Millipore), probed overnight with specific primary antibodies (α-tubulin: DM1A, Sigma T9026, 1:2,500), ((K(Ac)40)-alpha-tubulin: BML-SA452-0100, Enzo, 1:5,000), (histone H3: Abcam1791, 1:4,000), (acetyl-histone H3: Millipore 06-599, 1:4,000), followed by secondary antibodies in 5% non-fat dry milk in TBS blocking buffer. Detection was done with Amersham ECL western blotting reagent (GE Healthcare).

(*R*)- and (*S*)-TSA synthesis and purification. Racemic TSA was obtained via synthetic methods described in[58]. Racemic TSA was purified using a Waters XBridge Prep C18 5 µm OBD column (19 × 50 mm) and water/acetonitrile gradient (15–80% ACN in water, pH = 7; 18 min run, flow rate = 20 ml/min; racemic TSA RT = 8.0 min). With analytically pure racemic TSA in hand, a method was developed for the chiral separation. Racemic TSA was determined to be separable using a Daicel ChiralPAK AD-H 5 µm column (amylose tris(3,5-dimethylphenylcarbamate-coated 5 µm silica gel, 4.6 × 250 mm) and heptane/isopropanol gradient (10–60% IPA in heptane, no additives; 23 min run, flow rate = 3 ml/min). (*R*)-TSA eluted at RT = 15.81 min. (*S*)-TSA eluted at 14.41 min. These samples were collected and measured for their optical rotation. (*R*)-TSA $[a]^{20}_D$ = + 80.5 (*c* = 0.2, MeOH). (*S*)-TSA $[a]^{20}_D$ = −75.5 (*c* = 0.2, MeOH). ^{1}H NMR and LCMS data conform to previously reported structural data in ref. 39.

Cloning, expression and purification of mouse α-TAT. For expression of the α-TAT protein, a cDNA encoding mouse α-TAT (amino acids 1–197) was cloned into pOPINF vector via Gibson assembly (NEB) and expressed in Sf9 cells. Primers were constructed based on cDNA sequence (NP_001136216.1), mRNA was extracted from MEFs by using RNeasy mini kit (QIAGEN) as a template for cDNA. α-TAT was purified with the same procedures as described above for HDAC6 CD1. After first Ni-NTA IMAC agarose, His-tag was digested with 3C protease at 4 °C overnight. The cleaved α-TAT was further purified with second Ni-NTA agarose resin, then injected onto S200 gel filtration column (one step gel filtration). Purified α-TAT was assessed by *in vitro* acetylation assay with acetyl-CoA in the microtubules as previously described[17–19].

Tubulin and microtubule deacetylation assays. Microtubules were reconstituted with purified porcine brain tubulin or bovine brain tubulin (Cytoskeleton T238P or TL238, respectively). Microtubule reconstitution was done as described[59]; briefly, 10 mg/mL purified tubulin was polymerized with 2× Polymix (80 mM PIPES pH6.8, 1 mM EGTA, 1 mM $MgCl_2$, 2 mM GTP, 20% DMSO) for 40 min at 37 °C, then stabilized with pre-warmed BRB80-DT

(80 mM PIPES pH6.8, 1 mM EGTA, 1 mM $MgCl_2$, 5 mM DTT, 20 μM Taxol) buffer for 10 min. Taxol-stabilized microtubules were spun down 16,000*g* for 30 min at room temperature. Subsequently, polymerized microtubules were treated at 37 °C for 1 h with 5 μM mouse TAT (α-TAT) in the presence of ^{3}H-acetyl-CoA (0.1mCi/mL). Fully-acetylated microtubules were spun down at 16,000*g* for 30 min at room temperature, and washed with BRB80-DT buffer three times to remove α-tubulin α-TAT and unincorporated radioactivity of acetyl-CoA. Free tubulin dimers were generated from this material by cold treatment together with nocodazole: radiolabeled microtubules were spun down and dissolved in BRB80 buffer containing 5 mM DTT and 1 mM nocodazole. After incubation on ice for 1 h, residual MTs were spun down at maximum speed for 30 min. The supernatant was used as the free tubulin fraction. The absence of contamination with residual microtubules was checked by microscopy. For experiments with polymerized microtubules, MTs were stabilized with Taxol in BRB80-DT buffer. For the assays, different amounts of purified HDAC6 proteins were incubated at 37 °C for 1 h with equal amounts of radiolabeled acetylated microtubules or free tubulin. The reaction was stopped by adding 5% TCA, and the precipitated material was collected by a filter binding assay. The amount of precipitable ^{3}[H] tritium remaining in the tubulin or MTs was measured by liquid scintillation counting (Beckman). The level of tubulin deacetylase activity was determined by subtracting the counts obtained with HDAC6 treatment from those without HDAC6 treatment (i.e., input radioactivity).

Single molecule assay by immunofluorescence microscopy for microtubule deacetylation. Microtubules were reconstituted *in vitro* as mentioned above, and fully acetylated with 5 μM α-TAT and 250 μM acetyl-CoA, because tubulin from brain is not fully acetylated[60]. Deacetylation assays were performed with purified catalytic domains of zebrafish HDAC6 at 37 °C for different time point. Reactions were stopped by fixation with 1% glutaraldehyde, and dropped onto poly-Lys-coated coverslips for 10 min. Microtubules were stained with rabbit anti-AcK40 antibodies (BML-SA452-0100, Enzo, 1:400 dilution) and anti-beta tubulin antibodies (Sigma-Aldrich, 1:400 dilution), then detected by staining with Alexa-488 goat anti-rabbit (Lifetechnologies, 1:1,000 dilution) and Alexa-568 goat anti-mouse secondary antibodies (Lifetechnologies, 1:500 dilution). Microtubule images were captured with an Axioimager Z1 microscope (Zeiss) using a 100x objective lens. The acetylated microtubule signal intensity was traced and quantified by ImageJ (NIH).

48. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
49. Vonrhein, C. *et al.* Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 293–302 (2011).
50. McCoy, A.J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
51. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1002–1011 (2006).
52. Afonine, P.V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
53. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
54. Bricogne, G. *et al.* BUSTER version 2.11.4. (Global Phasing Ltd., Cambridge, UK, 2011).
55. Schrödinger, LLC. *The PyMOL Molecular Graphics System* Version 1.7.6 (2010).
56. Söding, J., Biegert, A. & Lupas, A.N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244–W248 (2005).
57. Martí-Renom, M.A. *et al.* Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 291–325 (2000).
58. Cosner, C.C. & Helquist, P. Concise, convergent syntheses of (±)-trichostatin A utilizing a Pd-catalyzed ketone enolate α-alkenylation reaction. *Org. Lett.* **13**, 3564–3567 (2011).
59. Ziółkowska, N.E. & Roll-Mecak, A. In vitro microtubule severing assays. *Methods Mol. Biol.* **1046**, 323–334 (2013).
60. Carbajal, A., Chesta, M.E., Bisig, C.G. & Arce, C.A. A novel method for purification of polymerizable tubulin with a high content of the acetylated isotype. *Biochem. J.* **449**, 643–648 (2013).

2.3 Appendix

2.3.1 Functional importance of two α-helices for deacetylation of α-tubulin by HDAC6 *in vivo*

Appendix 1

The structure of HDAC6 solved by our work (**2.2**) indicates that catalytic domains contain unique α-helices among HDAC family member proteins: Helix 6 and Helix 25 (in CD1 and in CD2 respectively). As only HDAC6 can deacetylate α-tubulin among HDAC family proteins, we hypothesized that these unique two α-helices are key molecular features for α-tubulin deacetylation. In the published work (**2.2**), the functional importance of these α-helices were tested *in vitro* using cellular lysate containing Ac-α-tubulin and purified HDAC6 catalytic domains, where Helix 6 and Helix 25 are replaced with a loop structure originating from HDAC8, respectively. In subsequent work, each α-helix was systematically replaced with the other α-helix or a HDAC8 loop, and the α-tubulin deacetylation activity was tested in HDAC6 rescuant MEFs. We found that it is necessary for CD2 to have an appropriate helix at the place where Helix 25 resides for α-tubulin deacetylation, and that Helix 6 is also important for deacetylation activities for α-tubulin, although this effect was not observed *in vitro* (**2.2**).

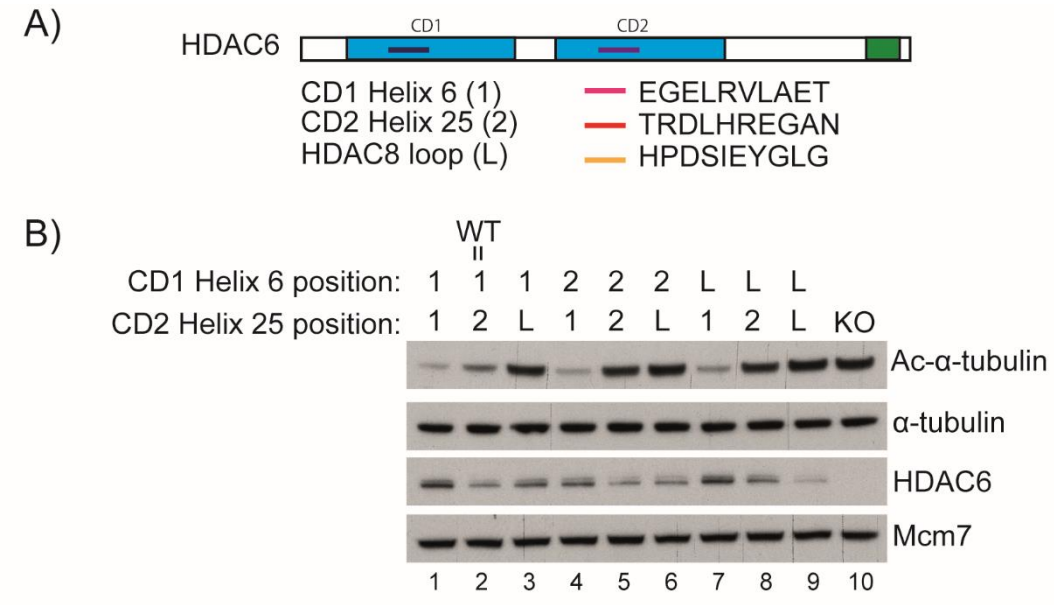


Figure legend: The functional importance of two helices for deacetylation of α-tubulin by HDAC6 (mouse) *in vivo*. (A) Sequences of HDAC6 specific α-helices and the loop derived from corresponding region of HDAC8. (B) Each HDAC6 WT and helix-mutant construct in retroviral pMSCV-IRES-EGFP vector was expressed in HDAC6 KO cells. EGFP-positive cell populations were sorted by FACS, and Ac-α-tubulin in each pooled population was tested by western blotting. Mcm7 was also tested as loading control.

2.3.2 Functional importance of N-terminal domain for deacetylation of α-tubulin by HDAC6 *in vivo*

Appendix 2

The laboratory of our collaborator, Dr. Cyril Barinka (Institute of Biotechnology CAS, Prague), found that the clusters of lysine(K)-rich positively charged residues in N-terminal domain of HDAC6 are necessary and sufficient for its binding to microtubule *in vitro*. Furthermore, they assume that these regions are acetylated, and that this charge neutralization impairs the binding ability. According to this finding, we hypothesized that the regions are also important for α-tubulin deacetylation activity of HDAC6. In order to test this hypothesis, catalytic activity of each HDAC6 N-terminal domain mutant was tested in HDAC6 KO 293T cells, which I had previously established by CRISPR/Cas9 genome editing (see **Fig. S7 in 2.1**). We revealed that both K-rich regions (KR1 and KR2) are important for the catalytic activity on α-tubulin, and also confirmed that positively charged lysines are necessary for the KR2 function.

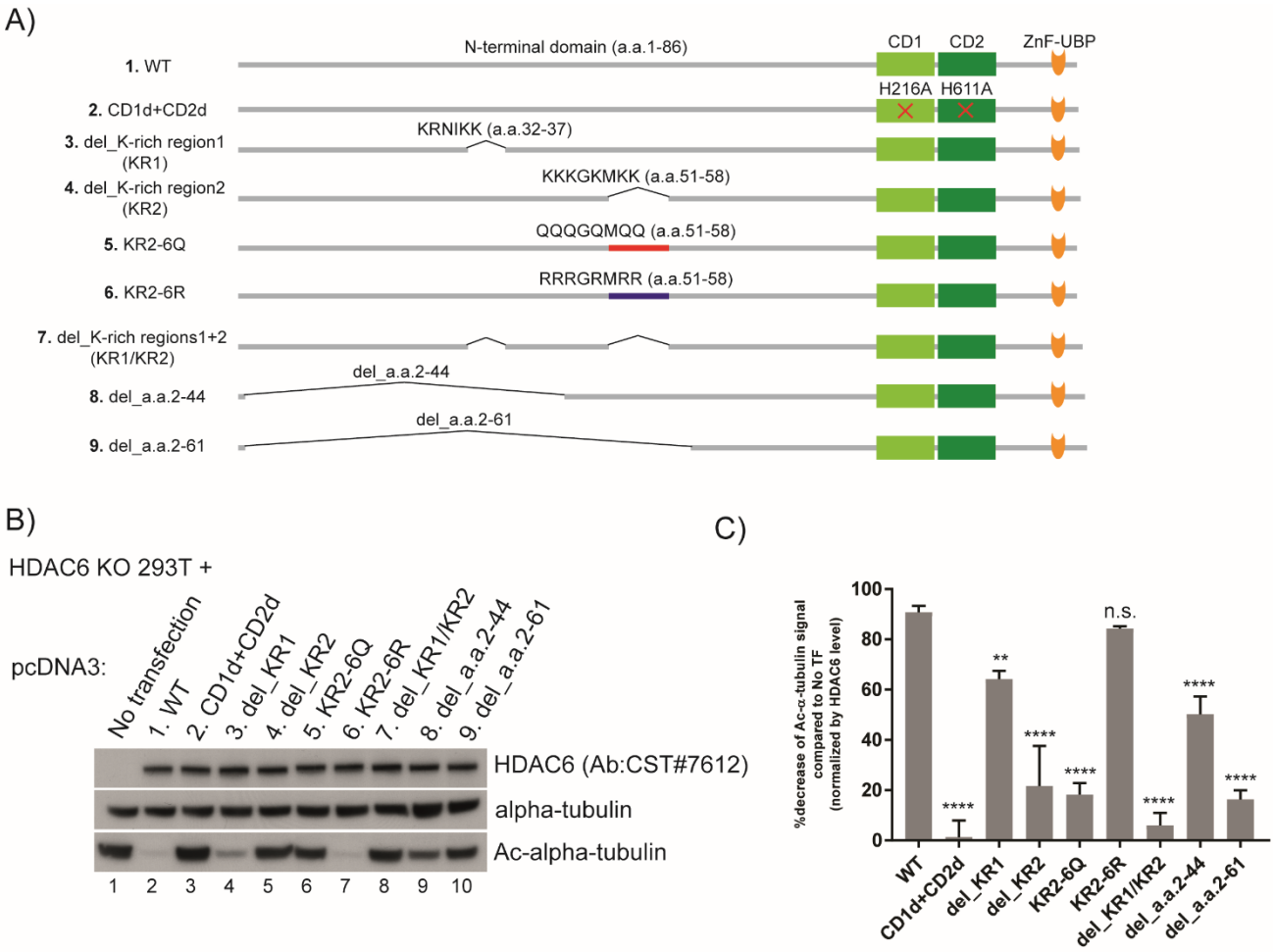


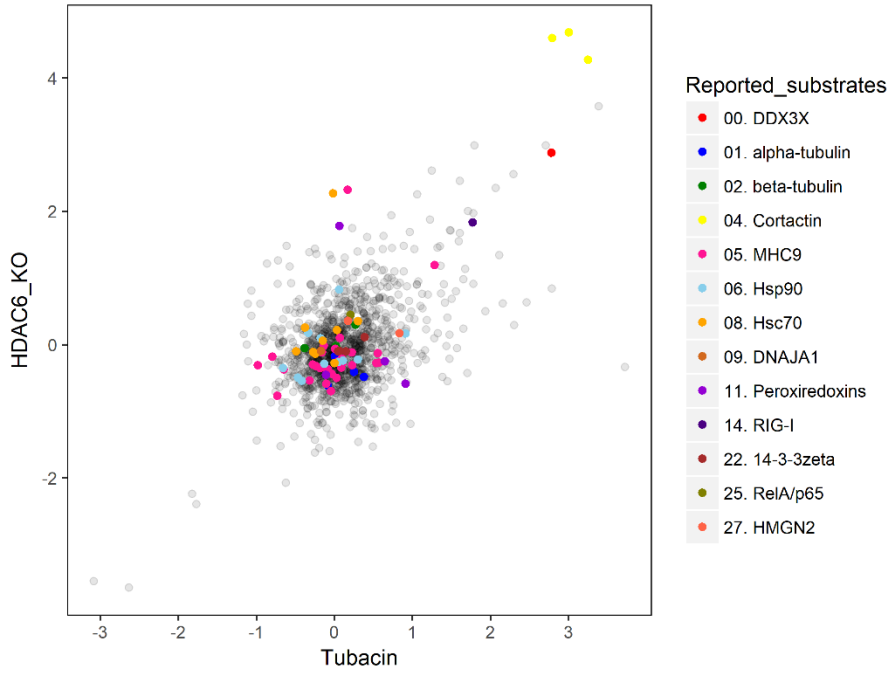
Figure legend: The functional importance of N-terminal domain for deacetylation of α-tubulin by HDAC6 (human) *in vivo*. (A) HDAC6 N-terminal domain deletion mutants used for the assay. HDAC6 harbors two K-rich sequences (K-rich region 1 and 2: a.a.32-37 and a.a.51-58, respectively) in its N-terminal domain (a.a. 1-86). Mutant 5 is acetyl-mimic mutant of KR2, while mutant 6 is acetyl-dead mutant. Mutants 8 and 9 are human constructs corresponding to zebrafish HDAC6 constructs, a.a. 25-831 and a.a. 40-831, respectively. (B) Each HDAC6 WT and mutant construct in pcDNA3 vector was transiently expressed in HDAC6 KO 293T cells, and Ac-α-tubulin level was tested by immunoblotting. (C) Quantification of Ac-α-tubulin signal in (B). The result is normalized by expression levels of α-tubulin and HDAC6. (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$ and (****) $P < 0.0001$ by one-way ANOVA.

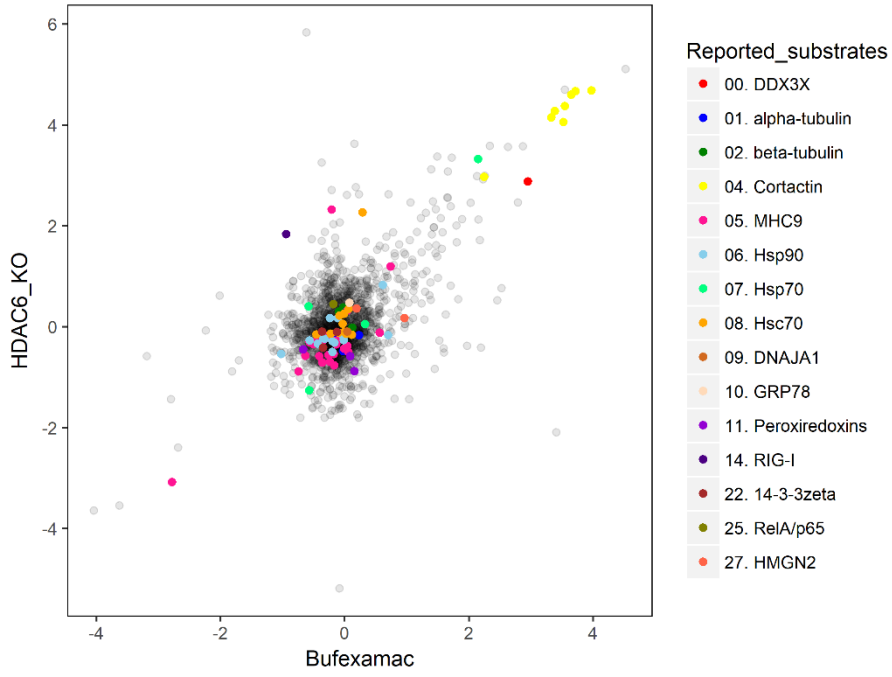
2.3.3 Previously reported HDAC6 substrates in the HDAC6-dependent acetylome

Appendix 3

HDAC6-dependent acetylome analysis identified 2029 acetylation peptides in total (**2.1**). To test the validity of the methodology, we determined which of the previously reported HDAC6 substrates (**1.2.2**) could be identified in our HDAC6-specific acetylome.

A)





B)

Reported substrates	Gene names	Position	Tubacin	Bufexamac	HDAC6_KO	VSL2	VL3-BA
00. DDX3X	Ddx3x;D1Pas1	118	2.78	2.95	2.87	D	D
01. alpha-tubulin	Tuba1b;Tuba4a;Tuba1a;Tuba1c;Tuba3a;Tuba8	96	0.25	0.01	-0.42	D	-
01. alpha-tubulin	Tuba1b;Tuba1a;Tuba1c;Tuba3a	326	0.01	-0.15	-0.19	-	-
01. alpha-tubulin	Tuba1b;Tuba4a;Tuba1a;Tuba1c;Tuba3a;Tuba8	370	0.38	-0.02	-0.49	-	-
01. alpha-tubulin	Tuba1b;Tuba4a;Tuba1a;Tuba1c;Tuba3a;Tuba8;Tubal3	394	0.01	0.23	-0.17	-	-
01. alpha-tubulin	Tuba1b;Tuba4a;Tuba1a;Tuba1c;Tuba3a;Tuba8	401	-0.08	N.D.	-0.61	-	-
02. beta-tubulin	Tubb2a	58	N.D.	-0.02	0.38	-	-
02. beta-tubulin	Tubb4b	58	0.27	-0.09	0.30	-	-
02. beta-tubulin	Tubb5	58	0.02	0.12	-0.02	-	-
02. beta-tubulin	Tubb3	58	-0.38	N.D.	-0.06	-	-
04. Cortactin	Cttn	87	3.00	3.98	4.68	D	D
04. Cortactin	Cttn	124	N.D.	3.52	4.06	D	D
04. Cortactin	Cttn	144	N.D.	2.25	2.97	D	D
04. Cortactin	Cttn	161	N.D.	3.55	4.37	D	D
04. Cortactin	Cttn	198	3.25	3.39	4.27	D	D
04. Cortactin	Cttn	235	N.D.	3.33	4.15	D	D
04. Cortactin	Cttn	272	N.D.	3.72	4.67	D	D
04. Cortactin	Cttn	309	2.79	3.65	4.60	D	D
05. MHC9	Myh9	8	1.29	0.74	1.19	-	-
05. MHC9	Myh9	74	-0.04	-0.24	-0.70	D	D
05. MHC9	Myh9	102	0.23	-0.13	-0.31	-	-
05. MHC9	Myh9	299	-0.18	-0.13	-0.35	-	-
05. MHC9	Myh9	355	0.03	0.06	-0.50	-	-
05. MHC9	Myh9	435	-0.08	N.D.	-0.37	-	-
05. MHC9	Myh9	439	N.D.	-2.78	-3.08	D	-
05. MHC9	Myh9	545	0.54	-0.13	-0.28	D	-
05. MHC9	Myh9	555	0.10	-0.25	-0.21	D	-
05. MHC9	Myh9	613	0.17	-0.21	2.32	-	-
05. MHC9	Myh9	856	-0.79	-0.29	-0.18	D	D
05. MHC9	Myh9	972	N.D.	-0.19	-0.28	D	D
05. MHC9	Myh9	1048	0.09	-0.51	-0.35	D	D
05. MHC9	Myh9	1181	-0.03	-0.01	-0.44	D	D
05. MHC9	Myh9	1209	0.23	0.56	-0.12	D	D
05. MHC9	Myh9	1249	0.07	N.D.	0.10	D	D
05. MHC9	Myh9	1330	0.57	N.D.	-0.27	D	D
05. MHC9	Myh9	1392	0.56	-0.07	-0.13	D	D
05. MHC9	Myh9	1404	0.01	N.D.	-0.07	D	D
05. MHC9	Myh9	1410	-0.13	N.D.	-0.02	D	D
05. MHC9	Myh9	1445	N.D.	-0.74	-0.89	D	D
05. MHC9	Myh9	1454	-0.10	-0.41	-0.59	D	D
05. MHC9	Myh9	1459	-0.98	-0.57	-0.31	D	D
05. MHC9	Myh11;Myh10;Myh9	1477	-0.73	-0.17	-0.77	D	D
05. MHC9	Myh9	1492	-0.16	0.02	-0.12	D	D
05. MHC9	Myh9	1513	N.D.	0.05	0.32	D	D
05. MHC9	Myh9	1525	-0.32	-0.17	-0.54	D	D
05. MHC9	Myh9	1614	0.04	-0.34	-0.26	D	D
05. MHC9	Myh9	1638	-0.14	-0.44	-0.34	D	D
05. MHC9	Myh9	1648	-0.65	0.04	-0.38	D	D
05. MHC9	Myh9	1669	0.11	-0.41	-0.28	D	D
05. MHC9	Myh9	1724	N.D.	-0.26	-0.56	D	D
05. MHC9	Myh9	1775	N.D.	-0.63	-0.58	D	D
05. MHC9	Myh9	1793	-0.28	-0.48	-0.30	D	D
05. MHC9	Myh9	1828	N.D.	-0.36	-0.71	D	D
05. MHC9	Myh9	1845	-0.24	N.D.	-0.32	D	D

Reported substrates	Gene names	Position	Tubacin	Bufexamac	HDAC6_KO	VSL2	VL3-BA
06. Hsp90	Hsp90ab1	53	-0.66	N.D.	-0.36	D	-
06. Hsp90	Hsp90aa1;Hsp90ab1	275	0.06	0.62	0.83	D	D
06. Hsp90	Hsp90aa1;Hsp90ab1	284	0.05	-0.01	-0.27	D	D
06. Hsp90	Hsp90aa1;Hsp90ab1	399	N.D.	0.70	-0.17	-	-
06. Hsp90	Hsp90ab1	435	0.30	-0.34	-0.23	D	-
06. Hsp90	Hsp90aa1	444	-0.17	N.D.	0.09	-	-
06. Hsp90	Hsp90ab1	481	N.D.	-0.43	-0.34	D	D
06. Hsp90	Hsp90aa1	490	N.D.	-0.16	-0.32	-	-
06. Hsp90	Hsp90ab1	531	0.11	-0.32	-0.24	D	D
06. Hsp90	Hsp90ab1	538	N.D.	-0.03	-0.17	D	D
06. Hsp90	Hsp90ab1	550	-0.13	-0.20	-0.29	D	D
06. Hsp90	Hsp90ab1	559	N.D.	-0.56	-0.27	D	D
06. Hsp90	Hsp90aa1	559	-0.13	-0.20	-0.29	D	D
06. Hsp90	Hsp90ab1	565	-0.47	-0.19	-0.50	D	D
06. Hsp90	Hsp90ab1	568	-0.42	-1.02	-0.55	D	D
06. Hsp90	Hsp90aa1	577	-0.33	-0.11	0.17	D	-
06. Hsp90	Hsp90b1	597	0.91	-0.24	0.17	D	D
07. Hsp70	Hspa4	194	N.D.	-0.57	-1.27	-	-
07. Hsp70	Hspa4	272	N.D.	0.33	0.05	D	D
07. Hsp70	Hspa4	559	N.D.	2.15	3.32	D	D
07. Hsp70	Hspa4	681	N.D.	-0.58	0.40	D	D
08. Hsc70	Hspa8	56	-0.15	-0.03	0.06	-	-
08. Hsc70	Hspa8	71	N.D.	0.12	-0.16	-	-
08. Hsc70	Hspa8	108	N.D.	-0.46	-0.16	D	-
08. Hsc70	Hspa2;Hspa8	187	0.30	0.07	0.35	-	-
08. Hsc70	Hspa8	246	-0.49	0.00	-0.10	D	D
08. Hsc70	Hspa8	319	-0.02	0.29	2.27	-	-
08. Hsc70	Hspa2;Hspa8	348	0.00	N.D.	-0.28	-	-
08. Hsc70	Hspa2;Hspa8	451	N.D.	-0.10	-0.12	-	-
08. Hsc70	Hspa8	512	-0.38	-0.01	0.25	D	D
08. Hsc70	Hspa8	524	0.03	-0.08	0.22	D	D
08. Hsc70	Hspa8;Hspa8	589	N.D.	-0.22	-0.14	D	D
08. Hsc70	Hspa8;Hspa8	597	-0.25	0.03	-0.14	D	D
08. Hsc70	Hspa8	601	-0.27	0.02	-0.11	D	D
09. DNAJA1	Dnaja1;Dnaja4;Dnajc5	32	-0.17	0.05	-0.10	D	D
10. GRP78	Hspa5	548	N.D.	0.08	0.47	D	D
11. Peroxiredoxins	Prdx1	27	0.91	0.09	-0.59	-	-
11. Peroxiredoxins	Prdx1	35	N.D.	0.16	-0.88	-	-
11. Peroxiredoxins	Prdx5	79	0.07	N.D.	1.78	-	D
11. Peroxiredoxins	Prdx1	136	0.65	N.D.	-0.25	-	-
11. Peroxiredoxins	Prdx6	209	-0.11	-0.66	-0.45	D	D
14. RIG-I	Ddx58	852	1.77	-0.94	1.83	D	-
22. 14-3-3zeta	Ywhaz	3	N.D.	-0.35	-0.41	D	D
22. 14-3-3zeta	Ywhaz	68	0.05	-0.12	-0.11	D	D
22. 14-3-3zeta	Ywhaz	115	0.07	N.D.	-0.10	-	-
22. 14-3-3zeta	Ywhaz	120	0.15	-0.36	-0.10	-	-
22. 14-3-3zeta	Ywhaz	157	0.39	N.D.	0.11	D	-
25. RelA/p65	Rela	310	0.21	-0.18	0.44	D	D
27. HMGN2	Hmgn2;Gm6724;Gm16494	73	0.84	0.97	0.17	D	D
27. HMGN2	Hmgn2	86	0.17	0.19	0.36	D	D

Figure legend: Previously reported HDAC6 substrates identified in the HDAC6-dependent acetylome. (A) The scatter plots show the correlation between acetylated sites quantified in HDAC6-specific inhibitor-treated MEFs vs acetylated sites quantified in HDAC6 KO MEFs. X-axis: $\log_2$ SILAC ratio of Tubacin-treated MEFs vs control-treated MEFs (Tubacin, upper panel) or $\log_2$ SILAC ratio of Bufexamac-treated MEFs vs control-treated MEFs (Bufexamac, bottom panel). Y-axis, both panels:

log_2 SILAC ratio of HDAC6 KO MEFs vs WT MEFs (HDAC6_KO, both panels). In each plot, the colored dots show the peptides originated from reported HDAC6 substrates (numbering of each substrate is based on the definition in **1.2.2**). As mentioned in **2.1**, DDX3X is the substrate identified by our study. (B) The log_2 SILAC ratio of identified acetylation sites originated from the 14 indicated proteins are listed. Positions written in red demonstrate acetylation sites previously reported as HDAC6 target sites. The columns, “VSL2” and “VL3-BA” show the disorder status of the identified site, which is defined by the VSL2 or VL3-BA programs of the PONDR suite (Prediction Of Naturally Disordered Regions). “D” indicates that the site is in IDRs while “-” indicates that the site is not in IDRs. N.D., not detected.

2.3.4 Quality check of commercial acetylated lysine antibodies

Appendix 4

First of all, evaluation of DDX3X acetylation in immunoblotting was required to further analyze its biological functions. Agarose conjugated acetyl-lysine antibody supplied from Immunechem (ICP0388) had been used for acetylome analysis, however it is not suitable for immunoblotting. Thus, in order to detect acetylation of DDX3X and other high-confidence HDAC6 target sites by immunoblotting, six commercial pan-acetyl-lysine antibodies were compared in parallel. We found that the CST (#9441) antibody exhibits a broad range of reactivity, and it was used in subsequent experiments (**2.1**).

A)

Name and supplier	Source	Immunogen
Immunechem, ICP0380, Anti Acetyl Lysine	Rabbit polyclonal	Acetylated KLH conjugates
Cell Signaling Technology #9441, Acetylated-Lysine	Rabbit polyclonal	A synthetic acetylated lysine-containing peptide
Cell Signaling Technology #9681, Acetylated-Lysine (Ac-K-103)	Mouse IgG2a	A synthetic acetylated lysine-containing peptide
Cell Signaling Technology #9814, Acetylated-Lysine (Ac-K2-100)	Rabbit IgG (mixture of 2 clones)	A synthetic acetyl-lysine peptide library
Millipore 05-515, Anti-Acetyl-Lysine Antibody (4G12)	Mouse IgG	Mixture of chemically acetylated antigens
Thermo Scientific, MA1-2021 Acetyl Lysine Antibody (1C6)	Mouse IgG	Synthetic peptide sequence surrounding H3K9Ac

B)

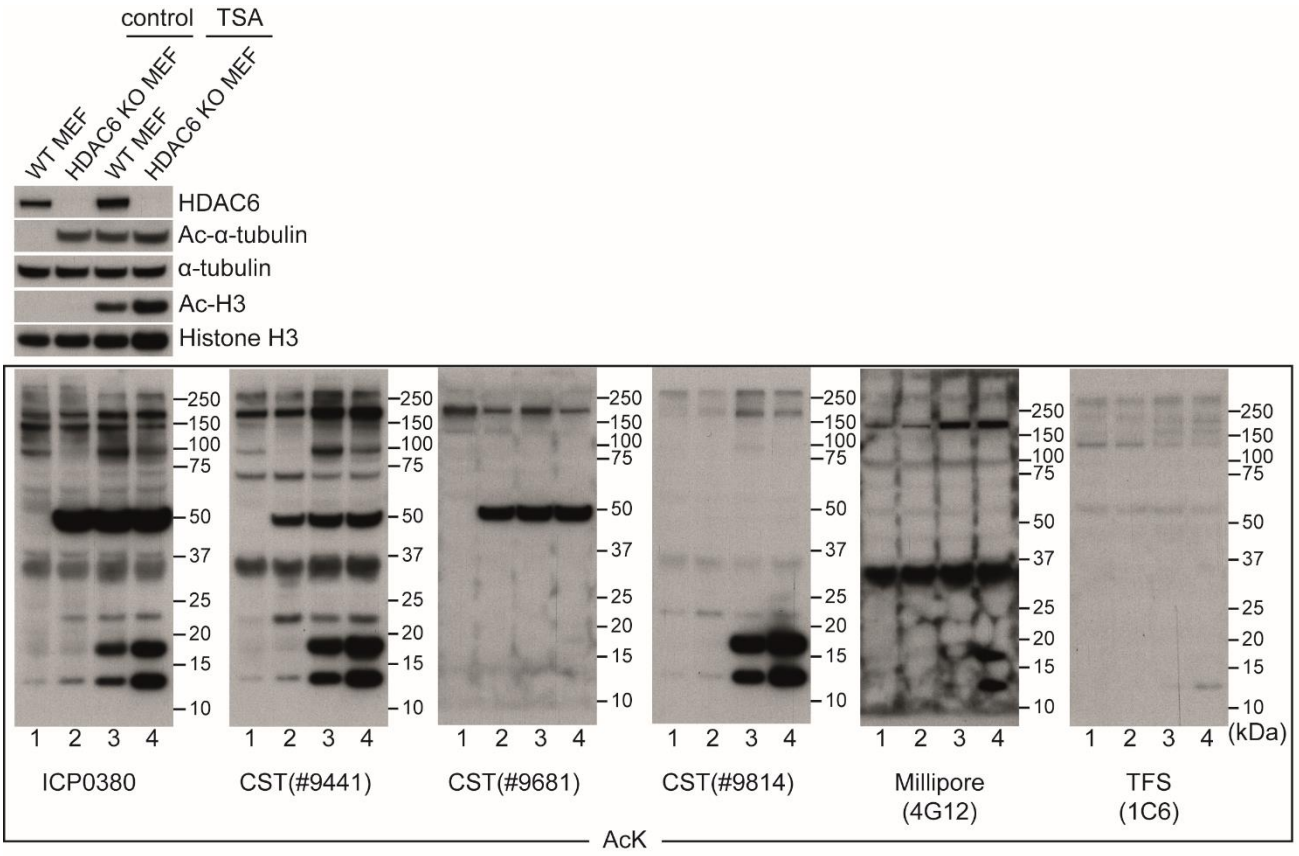


Figure legend: Quality check of commercial acetylated lysine antibodies. (A) Information about sources and immunogens of the antibodies used in (B). (B) WT and HDAC6 KO MEFs were treated with 2 μM TSA for overnight, so that the acetylation level of the proteins in these cells is upregulated. Then, lysates of these cells were immunoblotted with six types of acetylated lysine antibodies (ICP0380, CST#9441, CST#9681, CST#9814, Millipore (4G12) and Thermo Fischer Scientific (TFS) 1C6).

2.3.5 Pex5 is a novel HDAC6 substrate

Appendix 5

Among high-confidence HDAC6 target sites, the biological role of DDX3X acetylation was mainly studied as mentioned above (**2.1**). Another candidate, Pex5 (acetylation site: K28) is the peroxisomal import receptor, which binds to its cargos in the cytosol and brings them to translocation machinery at the peroxisomal membrane. Given that HDAC6 regulates cellular redox status through deacetylation of peroxiredoxins (**1.2.2**), it may also control the function of peroxisomes, which play important roles in regulating redox signaling pathways, through Pex5 deacetylation. To further explore this possibility, the Pex5 acetylation level under Tubacin was tested, and it was confirmed that Pex5 is a novel HDAC6 substrate.

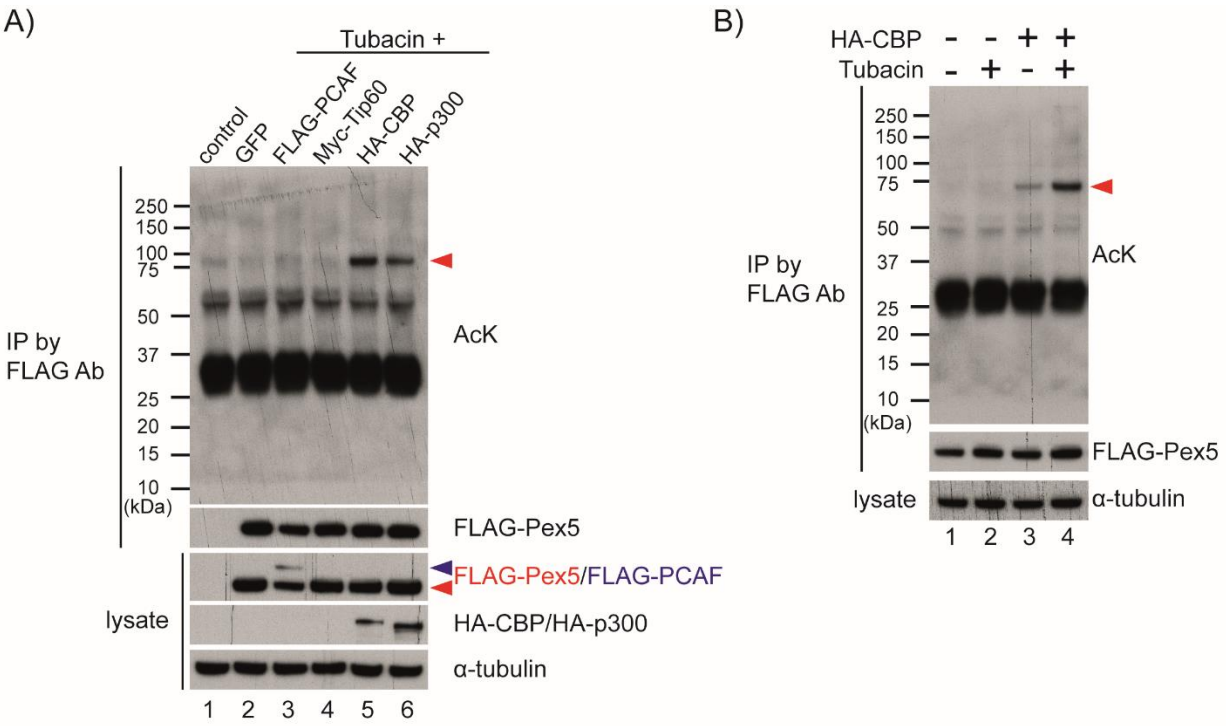


Figure legend: Pex5 is a novel HDAC6 substrate. (A) Acetylation of exogenous Pex5 by CBP and p300. Tagged versions of PCAF, Tip60, CBP and p300 were transiently transfected into 293T cells with FLAG-Pex5 (mouse). The acetylation status of exogenous Pex5 under Tubacin was analyzed by immunoblotting with a pan-acetyl-lysine antibody (AcK), following FLAG immunoprecipitation (IP). Cell lysate from parental 293T cells without Tubacin treatment was also loaded on the gel as a control

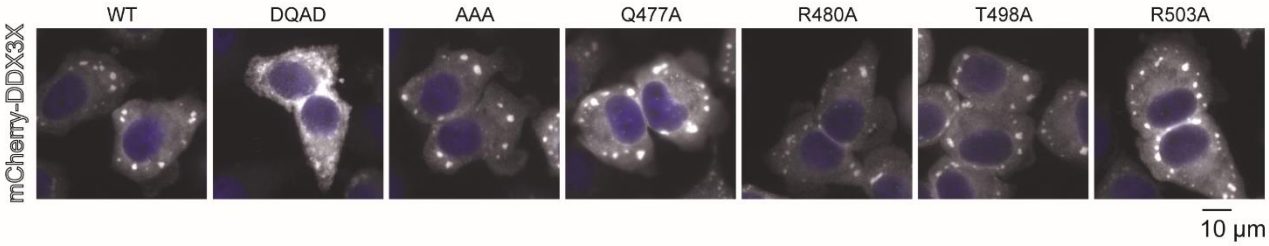
(lane 1). Red arrowhead indicates FLAG-Pex5 while blue one indicates FLAG-PCAF. (B) Acetylation of exogenous Pex5 is increased by Tubacin treatment. 293T cells were transiently transfected with FLAG-Pex5 together with CBP or GFP (control, -), and were treated overnight with 10 μM Tubacin, as indicated. Cell lysates were used to examine the acetylation status of FLAG-Pex5 by immunoblotting with a pan-acetyl-lysine antibody after immunoprecipitation. Red arrowhead indicates FLAG-Pex5.

2.3.6 SG formation ability of DDX3X catalytic activity mutants

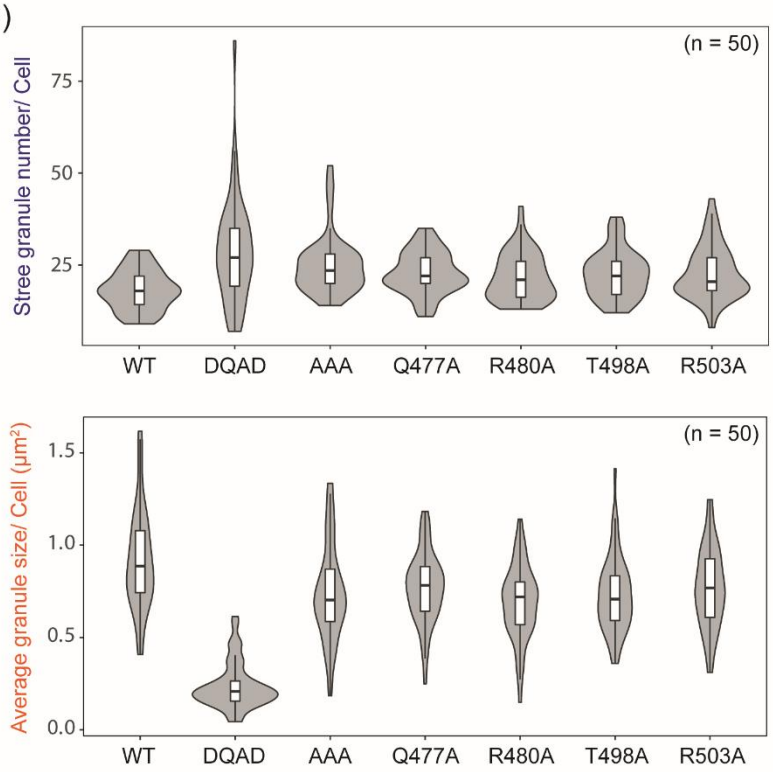
Appendix 6

In this thesis, it was shown that DDX3X contributes to SG formation through its IDR1 and helicase core domain (**Fig. S16 in 2.1**). As mentioned above, IDR1 undergoes LLPS, and thus leads to SG formation. On the other hand, the exact role of the DDX3X helicase core in SG formation was still under discussion. Its RNA binding ability contributes to SG formation, as also observed in other RNPs. At the same time, it appears that such ATP-dependent remodeling complexes including helicases and chaperones also support RNP granules disassembly (Protter and Parker, 2016). Thus, a systematic study to examine all three activities of the helicase core (RNA-binding, ATPase and RNA-unwinding) in SG formation is required. The molecular mechanisms coupling these three properties are still unknown, which makes it difficult to separately assess the contribution of each function in SG formation. However, a study about crystal structure of another DEAD-box helicase, Vasa (DDX4 ortholog of *D. melanogaster*), has provided some clues to uncouple these entangled functions (Sengoku et al., 2006). As DEAD-box helicase cores are well conserved between DDX3X and Vasa, functional key residues originally identified in Vasa were mutated in DDX3X, and SG formation ability of the DDX3X mutants were tested. We found that all six mutants formed smaller SGs, confirming that the helicase core is important for normal SG formation. Given that all mutants exhibit impaired RNA-unwinding ability, this primary helicase function may give a good explanation for these results.

A)



B)



Comparison	SG number	Average SG size
WT vs. DQAD	****	****
WT vs. AAA	***	***
WT vs. Q477A	ns	*
WT vs. R480A	ns	****
WT vs. T498A	ns	***
WT vs. R503A	ns	**
DQAD vs. AAA	ns	****
DQAD vs. Q477A	*	****
DQAD vs. R480A	***	****
DQAD vs. T498A	**	****
DQAD vs. R503A	*	****
AAA vs. Q477A	ns	ns
AAA vs. R480A	ns	ns
AAA vs. T498A	ns	ns
AAA vs. R503A	ns	ns
Q477A vs. R480A	ns	ns
Q477A vs. T498A	ns	ns
Q477A vs. R503A	ns	ns
R480A vs. T498A	ns	ns
R480A vs. R503A	ns	ns
T498A vs. R503A	ns	ns

C)

Vasa mutant	RNA Xlink (%)	ATPase (%)	Unwinding	Corresponding DDX3X mutant
WT	100	100	+++	WT
Q525A	61	70	-	Q477A
R528A	29	7	-	R480A
T546A	52	0	-	T498A
R551A	55	95	+	R503A

D)


sp|P09052|VASA1_DROME MSDDWDDEPIVDTRGARGGDWSDDEDTAKSFSGEAE---GDGVGGSGGEGGGYQGGNRDV 57
sp|Q62167|DDX3X_MOUSE MSH-------VAVENALG--------LDQQFAGLDLNSSDNQSGGSTASKGRYIPPHLR- 44
sp|P09052|VASA1_DROME FGRIGGGRGGGAGGYRGGNRDGGGFHGGRREGERDFRGGEGGFRGGQGGSRGGQGGSRGG 117
sp|Q62167|DDX3X_MOUSE ------NREATKGFY---DKDSSGWSSSKDKDA----------YSSFGSRGDSRGK 81
sp|P09052|VASA1_DROME QGGFRGGEGGFRGRLYENEDGDERRGRLDREERGGERRGRLDREERGGERGERGDGGFAR 177
sp|Q62167|DDX3X_MOUSE S-SFFG----------DRGSGSRGRFDDRGRG-------DYDGIGGRGDRSGFGKFER 121
sp|P09052|VASA1_DROME RRRNEDDINNNNNIVEDVERKREFYIPPEPSNDAIEIFSSGIASGIHFSKYNNIPVKVTG 237
sp|Q62167|DDX3X_MOUSE GG---------NSRWCDKSDEDDWSKPLPPSERLEQELFSGGNTGINFEKYDDIPVEATG 172
Q-Motif
Motif I
sp|P09052|VASA1_DROME SDVPQPIQHFTSADLRDIIIDNVNKSGYKIPTPIQKCSIPVISSGRDLMACAQTGSGKTA 297
sp|Q62167|DDX3X_MOUSE NNCPPHIESFSDVEMGEIIMGNIELTRYTRPTPVQKHAIPIIKEKRDLMACAQTGSGKTA 232
Motif Ia
sp|P09052|VASA1_DROME AFLLPILSKLLEDPHELE----------LGRPQVVIVSPTRELAIQIFNEARKFAFE 344
sp|Q62167|DDX3X_MOUSE AFLLPILSQIYADGPGEALRAMKENGRYGRRKQYPISLVLAPTRELAVQIYEEARKFSYR 292
Motif Ib
Motif Ic
Motif II
sp|P09052|VASA1_DROME SYLKIGIVYGGTSFRHQNECITRGCHVVIATPGRLLDFVDRTFITFEDTRFVVLDEADRM 404
sp|Q62167|DDX3X_MOUSE SRVRPCVVYGGAEIGQQIRDLERGCHLLVATPGRLVDMMERGKIGLDFCKYLVLDEADRM 352
Motif III
sp|P09052|VASA1_DROME LDMGFSEDMRRIMTHVTMRPE--HQTLMFSATFPEEIQRMAGEFLKNYVFVAIGIVGGAC 462
sp|Q62167|DDX3X_MOUSE LDMGFEPQIRRIVEQDTMPPKGVRHTMMFSATFPKEIQMLARDFLDEYIFLAVGRVGSTS 412
Motif IV
sp|P09052|VASA1_DROME SDVKQTIYEVNKYAKRSKLIEILSEQA--DGTIVFVETKRGADFLASFLSEKEFPTTSIH 520
sp|Q62167|DDX3X_MOUSE ENITQKVVWVEEIDKRSFLLDLLNATGKDSLTLVFVETKKGADSLEDFLYHEGYACTSIH 472
Motif IVa
Motif V
Motif Va
Motif VI
sp|P09052|VASA1_DROME GDRLQSQREQALRDFKNGSMKVLIATSVASRGLDIKNIKHVINYDMPSKIDDYVHRIGRT 580
sp|Q62167|DDX3X_MOUSE GDRSQRDREEALHQFRSGKSPILVATAVAARGLDISNVKHVINFDLPSDIEEYVHRIGRT 532
sp|P09052|VASA1_DROME GRVGNNGRATSFFDPEKDRAIAADLVKILEGSGQTVPDFLRTCGAGGD---------- 628
sp|Q62167|DDX3X_MOUSE GRVGNLGLATSFFNER-NINITKDLLDLLVEAKQEVPSWLENMAFEHHYKGSSRGRSKSS 591
sp|P09052|VASA1_DROME ---GGYSN--------------------------------QNFGGVDVRGRGN--YVGDA 651
sp|Q62167|DDX3X_MOUSE RFSGGFGARDYRQSSGASSSSFSSSRASSSRSGGGGHGGSRGFGGGGYGGFYNSDGYGGN 651
sp|P09052|VASA1_DROME TNVEEEEQWD- 661
sp|Q62167|DDX3X_MOUSE YNSQGVDWWGN 662

Figure legend: SG formation ability of DDX3X catalytic activity mutants. (A) Localization of DDX3X WT and its catalytic activity mutants transiently expressed in HeLa cells. Cells were treated with 1 mM arsenite for 30 min, localization of mCherry-DDX3X was assessed by microscopy after fixation. (B) Left: mCherry-DDX3X SG number (left panel) and size (middle panel) per cell were quantified (n = 50 cells in total, from biologically independent experiments) in (A), and displayed in violin plot; the thin line at the center represents the 95% confidence interval, and the bottom and top of the box represent the 25th and 75th percentile of the data, with the median value represented by a horizontal bold line. On each side of the thin line is a kernel density estimation showing the distribution shape of the data points from minima to maxima. Right: P values were determined by one-way ANOVA; (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$ and (****) $P < 0.0001$. (C) Functional status of Vasa mutants corresponding to DDX3X Q477A, R480A, T498A and R503A mutants used in (A) (cited from (Sengoku et al., 2006)). RNA Xlink (%) indicates the relative RNA-binding ability compared to that of WT. It is important to note that the functional status of both DQAD and AAA mutants was not assessed in their work. (D) Alignment of Drosophila Vasa and Mouse DDX3X. Conserved motifs in the DEAD-box and residues mutated in this experiment are highlighted (cyan: ATP binding, magenta: RNA binding, light green: Link ATP and RNA binding, and yellow: mutated residues).

2.3.7 LLPS efficiency of DDX3X-IDR1 with cancer-related mutations

Appendix 7

DDX3X can undergo LLPS, which is linked to its SG formation ability (**2.1**). DDX3X is recurrently mutated in some specific types of tumors, and the mutations are extensively mapped to its helicase core domain. These cancer-associated mutations in the helicase core domain drive SG formation and impair global translation (Valentin-Vega et al., 2016). Assuming from that certain oncogenes such as K-Ras enhance cell fitness through upregulated SG formation under stresses, this effect on DDX3X also contributes to tumor cell fitness. Furthermore, a few cancer-associated mutations in DDX3X-IDR1 have been registered in the COSMIC database (although there are fewer than the helicase core mutations). As I have shown that acetylation of a single lysine residue affects LLPS and SG formation (**Fig. 4 in 2.1**), it is possible that these cancer-associated point mutations in IDR1 also may affect these properties. To test this hypothesis, IDR1 harboring individual cancer-associated mutation were purified and their LLPS propensity was tested by droplet formation assay *in vitro*. As is visible, several of the cancer associated mutations, such as D67Y or R113G among others, show an altered LLPS propensity.

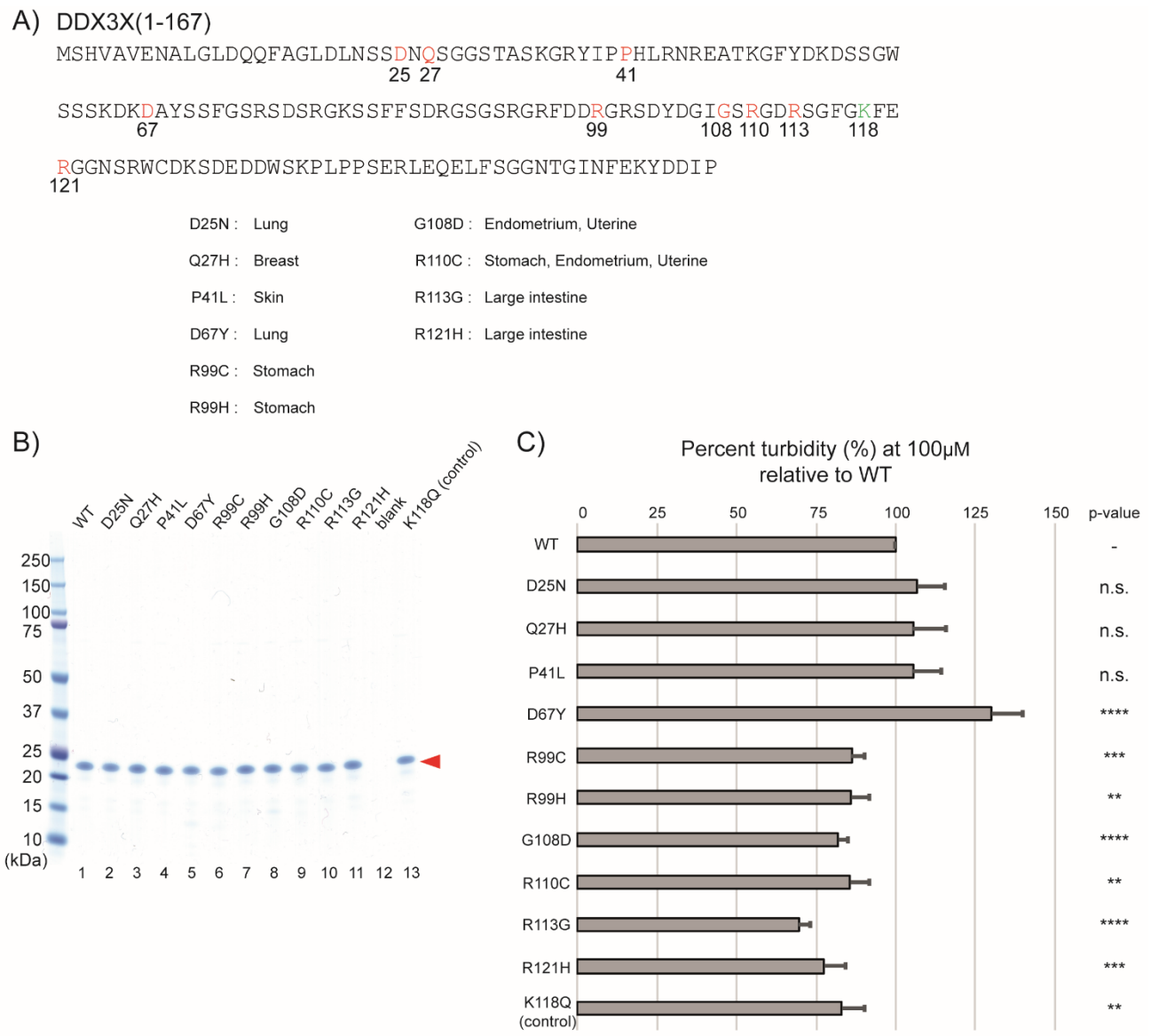
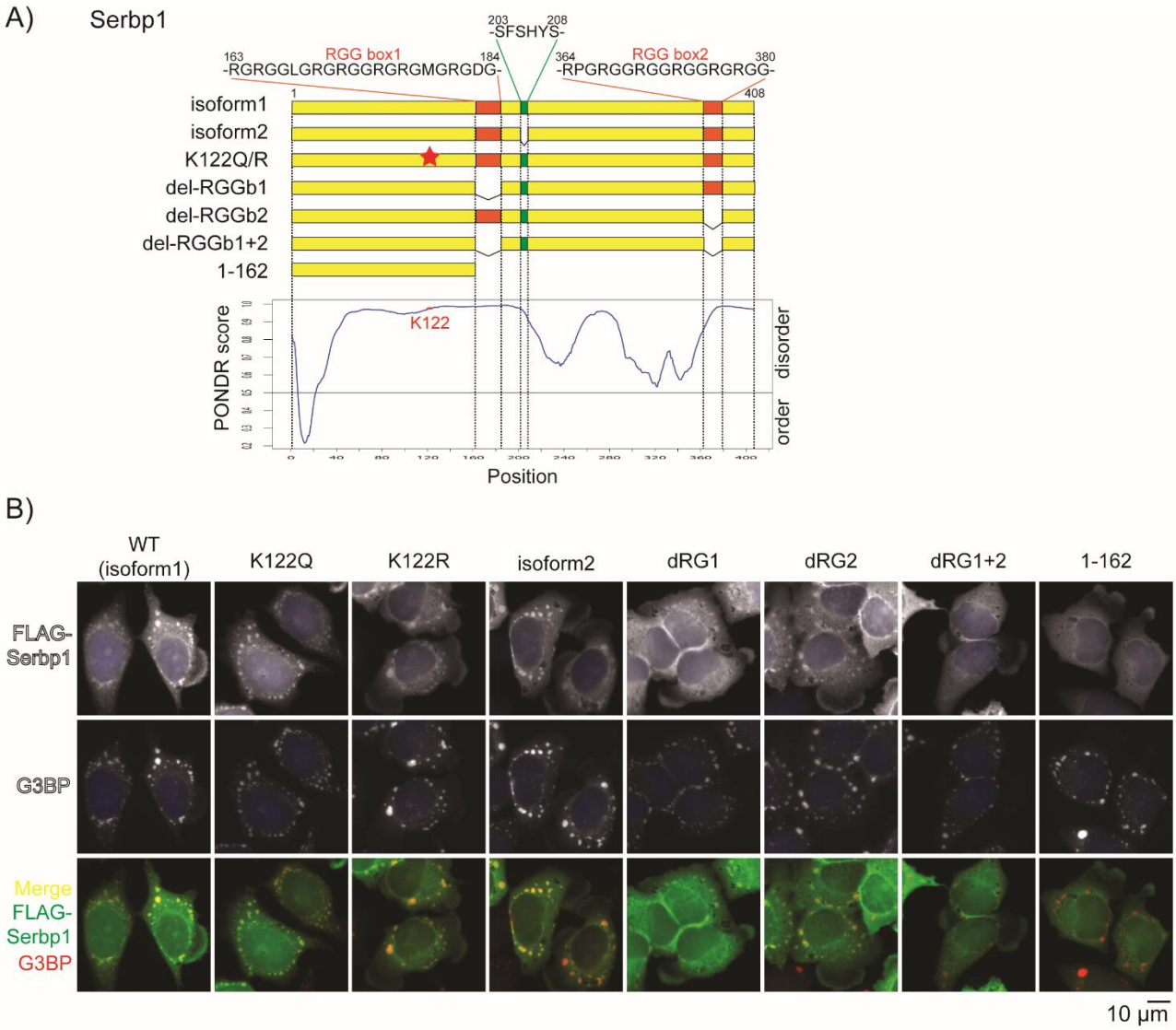


Figure legend: LLPS efficiency of DDX3X-IDR1 with cancer related mutations. (A) Mutations located in DDX3X-IDR1 found in cancer samples (registered in COSMIC https://cancer.sanger.ac.uk/cosmic, and cited in (Sharma and Jankowsky, 2014)). (B) A Coomassie Brilliant Blue stained SDS-PAGE gel showing the purified WT and mutants of DDX3X-IDR1 (red arrowhead). (C) Percent of turbidity (OD600) of DDX3X-IDR1 mutants normalized to WT (mean ± SD, n = 3). (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$ and (****) $P < 0.0001$ by Student's two-tailed t-test.

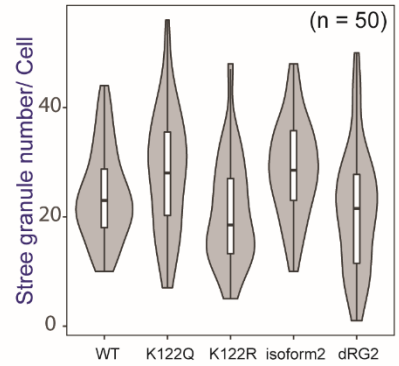
2.3.8 SG formation ability of Serbp1 acetyl-mimic/dead mutants

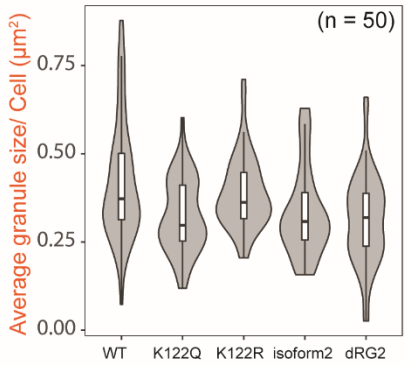
Appendix 8

HDAC6-dependent acetylome analysis revealed that a large part of high-confidence HDAC6 target sites map to IDRs (**Fig. 1 and S2 in 2.1**), suggesting that HDAC6 globally regulates LLPS propensity of its target proteins through their IDR deacetylation. In addition to DDX3X, another SG component protein Serbp1 (Lee et al., 2014b) is in the list of high-confidence HDAC6 target sites (K122 of Serbp1, **Fig. 1 and S2 in 2.1**). Thus, we hypothesized that HDAC6 also regulates SG formation of Serbp1 through its deacetylation. Serbp1, a RNA binding protein, harbors a long IDR containing two RGG/RG-rich sequences (RGG box1 and 2) (discussed in **1.3.2**), and also has an isoform missing 6 amino acids in the IDR. Thus, to test the importance of Serbp1 acetylation in SG formation, acetyl-mimic/dead mutant, RGG box deletion mutants as well as isoform-specific 6 amino acids deletion mutant were constructed and their SG formation ability was tested in HeLa cells. We confirmed that RGG box, which is often associated LLPS propensity, is also important for Serbp1 to localize on SGs. Moreover, isoform-specific short 6 amino acids also contribute to SG formation ability. As observed in DDX3X, acetyl-mimic mutation (K122Q) in Serbp1 IDR also affected SG formation; numerous smaller SGs were induced, and this effect was not observed with K122R mutant. Thus, HDAC6 may regulate the acetylation status of IDRs in a set of SG components.


A)
Serbp1
203
208
-SFSHYS-
163
RGG box1
184
364
RGG box2
380
-RGRGGLGRGRGGRGRGMGRGDG-
-RPGRGGRGGRGGRGRGG-
1
408
isoform1
isoform2
K122Q/R
del-RGGb1
del-RGGb2
del-RGGb1+2
1-162
K122
PONDR score
disorder
order
Position
B)
WT (isoform1)
K122Q
K122R
isoform2
dRG1
dRG2
dRG1+2
1-162
FLAG-Serbp1
G3BP
Merge
FLAG-Serbp1
G3BP
10 μm
C)
% of cells with Serbp1+ SGs in transfected cells
100
80
60
40
20
0
WT (isoform1)
K122Q
K122R
isoform2
dRG1
dRG2
dRG1+2
1-162

D)





Comparison	SG number	Average SG size
WT vs. K122Q	*	**
WT vs. K122R	ns	ns
WT vs. isoform2	*	**
WT vs. dRG2	ns	***
K122Q vs. K122R	**	*
K122Q vs. isoform2	ns	ns
K122Q vs. dRG2	**	ns
K122R vs. isoform2	***	*
K122R vs. dRG2	ns	*
isoform2 vs. dRG2	**	ns

Figure legend: SG formation ability of Serbp1 acetyl-mimic/dead mutants. (A) Serbp1 (mouse) domain deletion mutants and acetyl-mimic/dead mutants used for the assay. Serbp1 harbors two RGG/RG-rich sequences (RGG box1 and 2, orange), and the short insertion (a.a. 203-208: SFSHYS, green) is removed from isoform2 by alternative splicing. Residue K122 in its IDR is identified as one of high-confidence HDAC6 target sites (see **Fig. 1 and S2 in 2.1**). (B) Each FLAG-Serbp1 WT and mutant construct in pcDNA3 vector was transiently expressed in HeLa cells, and localization of FLAG-Serbp1 after arsenite treatment (1 mM, 30 min) was tested by IF. (C) The number of cells having Serbp1-positive SGs in (B) were counted, and percentage of all transfected cells are shown. (D) Quantification of Serbp1-positive SG number (left panel) or size (middle panel) per cell (n = 50 cells in total, from biologically independent experiments), and displayed in violin plot; the thin line at the center represents the 95% confidence interval, and the bottom and top of the box represent the 25th and 75th percentile of the data, with the median value represented by a horizontal bold line. On each side of the thin line is a kernel density estimation showing the distribution shape of the data points from minima to maxima. The five construct which induced over 50% SG positive cells in (B) were specifically tested. P values were determined by one-way ANOVA (right panel); (*) $P < 0.05$, (**) $P < 0.01$ and (***) $P < 0.001$.

Discussion

3.1 Novel mechanical insights for HDAC6-mediated deacetylation

3.1.1 Deacetylation mechanism by HDAC6

The crystal structure of HDAC6 catalytic domains was recently solved by two groups including our group (Hai and Christianson, 2016; Miyake et al., 2016) (see **2.2**). A prevalent model about the mechanism of HDAC6-mediated deacetylation suggested that only CD2 deacetylates its substrates including α-tubulin and cortactin *in vivo*. Our work successfully refined the model with providing important insights about both newly discovered HDAC6 specific α-helices and CD1.

Two α-helices for α-tubulin recognition

HDAC6 catalytic domains contain unique α-helices among HDAC family member proteins: Helix 6 and Helix 25 (in CD1 and in CD2 respectively). Replacement of CD2 Helix 25 by a loop structure coming from HDAC8 impairs its catalytic activity on α-tubulin, but not on acetylated lysine containing peptide *in vitro*. By contrast, replacement of CD1 Helix 6 by a loop structure originating from HDAC8 does not affect on its catalytic activity on α-tubulin *in vitro*. This suggests that Helix 25 is important to recognize α-tubulin for its deacetylation even *in vitro* context. Subsequent analysis with corresponding mouse HDAC6 α-helices mutants (**2.3.1 Appendix 1**) in MEF cells provided further insight about α-tubulin deacetylation. In agreement with *in vitro* α-tubulin deacetylation, replacement of CD2 Helix 25 by a loop structure impaired its α-tubulin deacetylation (**2.3.1 Apx. 1B; lanes 3, 6 and 9**). Interestingly, replacement of Helix 25 by Helix 6 did not impair its deacetylation activity (**2.3.1 Apx. 1B; lanes 1, 4 and 7**), suggesting that it is necessary for CD2 to have an appropriate helix at the place where Helix 25 resides for α-tubulin deacetylation. At the same time, it clearly demonstrated that HDAC6 constructs lacking Helix 6 (**2.3.1 Apx. 1B; lanes 5, 6, 8 and 9**) have impaired deacetylation activities for α-tubulin, although this effect was not observed *in vitro*. As far as we know, this is the first evidence to show the importance of CD1 domain for α-tubulin deacetylation; it seems that CD1 helps the process through this Helix 6, which fits well with previous observations from our laboratory (Zhang et al., 2006). We would like to propose a two-step deacetylation model for HDAC6-mediated α-tubulin deacetylation. First, Helix 6 in CD1 may recognize and bind to the α-tubulin or β-tubulin in

tubulin dimers. Second, catalytic pocket of the CD2 may contact with α-tubulin, where deacetylase activity of CD2 requires Helix 25 (or similar helix structure). The structure of HDAC6 binding to unpolymerized tubulin dimers by the X-ray crystallography or cryo-EM should be clarified in future to delineate the process completely.

In parallel, we also clarified that HDAC6 prefers to deacetylate unpolymerized tubulin, suggesting that it is acting as a resetting factor for acetylation status of tubulins once microtubule catastrophe takes place. By contrast, the initial work demonstrated purified HDAC6 potently deacetylated α-tubulin in assembled microtubules *in vitro* (Hubbert et al., 2002). This discrepancy may be due to the difference in assay system, HDAC6 protein species etc. and further investigations are required to determine the exact mechanism for α-tubulin K40 deacetylation by HDAC6 *in vivo*.

CD1 as a deacetylase for intrinsically disordered regions?

During the process to solve HDAC6 structure, we have also realized that the CD1 domain itself has catalytic activity, and deacetylates peptide substrates, but not α-tubulin *in vitro*. This observation has encouraged us to test CD1 catalytic activity for acetylated lysines located in IDRs, as we considered that IDRs may be regarded as long unstructured peptide *in vivo*. Expectedly, we observed that two of the high-confidence substrates whose acetylation site is in an IDR, DDX3X and cortactin, can be robustly deacetylated by CD1, as well as by CD2 (see **Fig. S8 in 2.1**). This is in contrast to deacetylation of α-tubulin, which strictly depends on CD2. These observations suggest that CD1 may be generally active on acetylated lysines in IDRs. It would be interesting to test whether or not mutation in Helix 6 and 25 affect HDAC6 catalytic activity on IDRs in future. This finding has important implications for the generation of HDAC6-specific inhibitors, which so far have been developed on the assumption that only CD2 is active, and inhibition of CD1 was usually not tested.

N-terminal domain for α-tubulin recognition

As mentioned so far, the molecular mechanism of α-tubulin deacetylation by HDAC6 is still under extensive investigation. Needless to say, its unique tandem deacetylase domains are central part of the study. In contrast, the N-terminal and C-terminal extensions including ZnF-UBP have not been

widely accepted as regulatory domains for α-tubulin deacetylation. The laboratory of our collaborator, Dr. Cyril Barinka, found that two clusters of lysine (K)-rich positively charged residues (KR1: a.a. 32-37 and KR2: a.a. 51-58) in N-terminal domain of human HDAC6 are necessary and sufficient for binding to microtubule *in vitro*. They also assumed that these regions are subjected to acetylation, and that the neutralization of positive charge impairs the binding ability. This hypothesis was partially supported by our observation that either KR1 or KR2 deletion impairs the catalytic activity on α-tubulin (**2.3.2 Apx. 2B; lanes 4 and 5, and 2C**). KR2 is likely to have more significant contribution to the activity, and furthermore, KR2 acetyl-mimic mutant (KR2-6Q) shows decreased activity on the same level with its deletion (**2.3.2 Apx. 2B; lanes 5 and 6, and 2C**). On the other hand, KR2 acetyl-dead mutant (KR2-6R) still has catalytic activity, which is comparable with that of WT (**2.3.2 Apx. 2B; lanes 2 and 7, and 2C**). Accordingly, it seems that the positive charge of KR2 (a.a. 51-58) may play an important role for α-tubulin deacetylation in cells (**2.3.2 Apx. 2**). Two N-terminal domain deletion mutants, which are the human version of the zebrafish constructs used for crystallization (**2.2**), also exhibit less activity, which perfectly fits with the observations presented here (**2.3.2 Apx. 2B; lanes 9 and 10, and 2C**). A few studies reported the importance of N-terminal region of HDAC6; this region contains a nuclear localization signal (NLS, a.a. 14-58) and a nuclear export signal (NES, a.a. 67-76) (Bertos et al., 2004). Given that the NLS including KR2 is important for HDAC6 to localize in the nucleus, KR2 mutant would prefer to stay in cytosol and actively deacetylate α-tubulin; however, KR2 could not deacetylate it efficiently. This discrepancy indicates that the effect of KR2 on α-tubulin deacetylation observed here is separate from its localization; obviously further studies are required to clarify the interactions between N-terminal region and tandem deacetylase domains of HDAC6.

3.2 HDAC6 regulates stress granule formation through DDX3X deacetylation

3.2.1 Identification of novel substrates of HATs/HDACs by acetylome analysis

In order to grasp a comprehensive picture of HDAC6 substrates, we utilized acetylome analysis combined with genetic and pharmacologic inhibition of HDAC6 (see **Fig. 1 and S1 in 2.1**). In total, approximately 2000 peptides were identified, and 1% of them were defined as high-confidence HDAC6 target sites that showed a more than 2-fold increase in acetylation in all tested conditions.

Our acetylome covers 103 acetylation sites from 14 reported substrates, including cortactin, a well-known substrate (**2.3.3 Apx. 3**). Most strikingly, eight cortactin acetylation sites were detected in three independent acetylomes (Tubastatin, Bufexamac, and HDAC6 KO), confirming all the four sites previously reported. Additionally, four novel acetylation sites were revealed. Conversely, the canonical HDAC6 target site α-tubulin K40 peptide was not detected, although five acetylated peptides from α-tubulin were identified. This may be due to lack of lysines and arginines, which are targeted by two digestion enzymes (Lys-C and trypsin) used in acetylome analysis in the AcK40 neighborhood. Consequently, the associated peptides, including AcK40, are too large to be detected by mass spectrometry. In addition to the identification of the four cortactin sites, the identification of β-tubulin K58, Hsp90 K284, and 14-3-3ζ K120 were consistent with the findings of previous studies. However, their acetylation levels do not seem to be upregulated neither by HDAC6 inhibitors nor in the KO condition ($\log_2$ SILAC ratio < 0.5). Acetylation sites of other substrates, such as MHC9 and Hsp70, have not been identified previously. Moreover, none of the acetylation sites from MHC9 and Hsp70 exhibit significant differences. The high-confidence HDAC6 target sites consist of an entirely new set of substrates, except cortactin. Differences in acetylation levels of the reported substrates have been determined using immunoblotting in previous studies. Considering that mass spectrometry is more sensitive compared to immunoblotting, it is tempting to think that acetylome could efficiently confirm significant changes in the acetylation status of previously reported substrates. To the best of our knowledge, this is the first study to revisit HDAC6 substrates utilizing acetylome analysis. Future studies are anticipated to resolve this discrepancy, with several possibilities (Fig. 15).

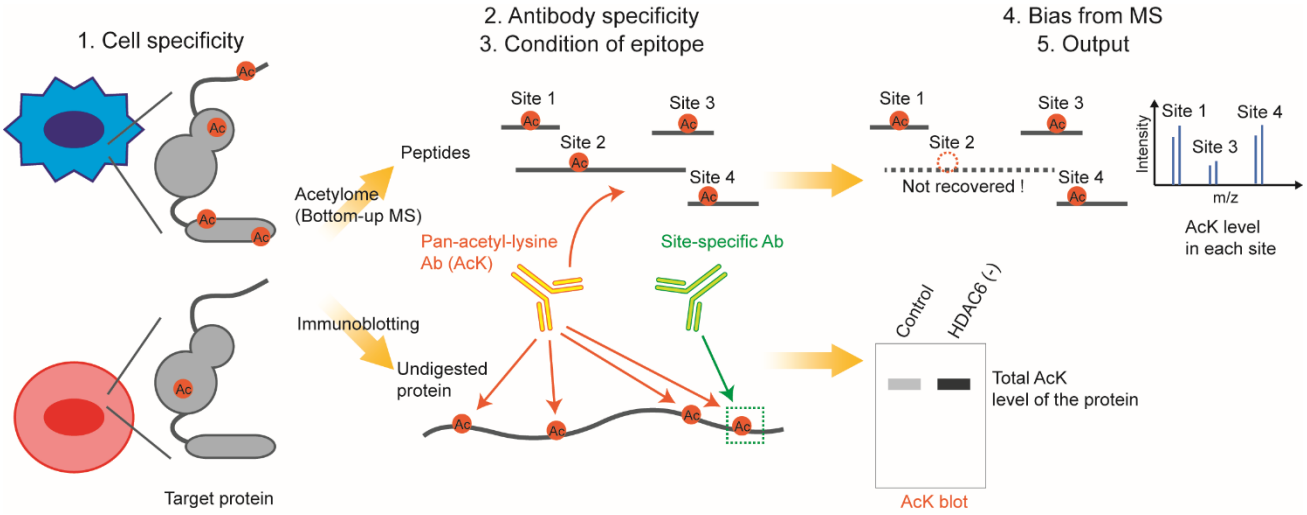


Figure 15. Comparison between acetylome and immunoblotting to confirm HDAC6 substrates. Five potential reasons to explain the discrepancy between previous work about HDAC6 substrates identification and our acetylome are presented. Acetylation status of HDAC6 substrate may depend on cell type (1. Cell specificity). Several commercial pan-acetyl-lysine (AcK) antibodies, and also site-specific antibodies are available, but there is no consensus about their quality in the field (2. Antibody specificity). The experimental condition for antibody-antigen reaction is different; digested peptides are immunoenriched with AcK antibodies in the acetylome analysis while undigested denatured proteins are reacted with antibodies on the immunoblot membrane (3. Condition of epitope). Relatively large peptides are not detected by mass spectrometry (4. Bias from MS). Acetylome analysis measures acetylation status of each lysine residue individually, whereas immunoblotting detects the acetylation status of the whole protein (5. Output).

Firstly, the repertoire of HDAC6 substrates may be greatly influenced by cell type specificity. We reported a HDAC6-dependent acetylome in MEFs, whereas previous studies employed different cells, including HeLa and other cancer cell lines to define HDAC6 substrates. Even if HDAC6 and other major substrates are ubiquitously expressed in almost all tissues, the regulatory system and how HDAC6 targets its substrates may be specific in every cell type. Secondly, the detection of protein acetylation in immunoblotting and acetylome analysis entirely depends on the affinity of antibodies, which could be another source of diverse contentious results. In some well-known acetylated proteins,

including histones and α-tubulin, highly site-specific antibodies are available. However, generating such antibodies is often time-consuming and uneconomical. Therefore, “pan”-acetyl-lysine antibodies are routinely applied in analyses. Moreover, to increase protein coverage, using a pan-acetyl-lysine antibody is more appropriate for acetylome analysis than a site-specific antibody for a specific protein. However, there is no consensus on “pan”-acetyl-lysine antibodies in acetylation research and their quality and specificity may vary (**2.3.4 Apx. 4**). Consequently, results vary substantially across studies. Further, even if the exact same antibody was to be utilized, the affinity would be altered based on the condition of its epitope. The antibody binds to its epitope on denatured protein in immunoblotting, whereas it binds to digested peptides in acetylome analysis. This makes it more challenging to compare results from two different experimental setups. Developing alternative universal approaches for detection (e.g. chemical probes directly recognizing acetyl-lysine residues without being affected by diverse physicochemical conditions) could benefit acetylation research in the future. Thirdly, bottom-up strategies currently utilized in acetylome analysis effectively provide information on each individual acetylation “site” in target proteins, which is not easily integrated into their whole acetylation “status” derived from co-occurrence of acetylation as detected by immunoblotting. There are often several acetylation sites as clearly observed in the reported substrates in our acetylome. Even if each acetylation site were weakly regulated by HDAC6, an assumption based on the minor changes in their SILAC ratios, the overall effects of several acetylations on one protein species could be robustly detected using immunoblotting with pan-acetyl-lysine antibodies. Although top-down strategies analyzing undigested proteins may resolve this problem and provide information about the co-occurrence of acetylation on the same protein, they are not yet suitable for proteome-scale analyses. This gap should be addressed to facilitate the understanding of the biology of acetylation (or other PTMs), because it seems that several acetylations often collectively alter biological function in target proteins, which is the case in IDR acetylation in cortactin and DDX3X. Finally, the interpretation of the acetylome data is influenced by the intrinsic bias associated with mass spectrometry analysis. As exemplified above by α-tubulin K40, the use of trypsin and Lys-C in an analysis limits acetylation site

identification based on their digestion capacities. To provide further protein coverage, additional alternative proteases, such as AspN and GluC, should be applied, although they are usually less specific and economical. Thus, it remains a major challenge to identify novel substrates of target PTM enzyme(s) utilizing conditional proteomes, where the PTM enzymes are inhibited or knocked out, and clarify their biological implications. Nevertheless, our study revealed important novel insights in the fields of protein acetylation, IDPs, and phase separation, beyond the biology of HDAC6 itself.

Analysis of these high-confidence HDAC6 target sites defined a HDAC6 target motif where several positions show an enrichment, in particular the two residues immediately before the acetylated lysine -GG**K(Ac)**-. Of note, this is reminiscent of the so-called GK motif associated with histone acetylation by CBP and p300 (Bannister et al., 2000). The well-established HDAC6 substrate, cortactin, also has this GK motif in its repeat region (-GXGG**K(Ac)**XG-) (Zhang et al., 2007). As mentioned in **1.2.2**, K58 on β-tubulin is often regarded as the equivalent site to K40 on α-tubulin because these sites are characteristically exposed to the luminal face of microtubules (Janke and Bulinski, 2011), and K58 is indeed deacetylated by HDAC6 (Liu et al., 2015). Surprisingly, K58 on β-tubulin fits well to this -GG**K(Ac)**- motif often observed in histone acetylation sites, while K40 on α-tubulin does not. This GK motif on β-tubulin seems to bridge between two functionally unrelated proteins, histones and tubulins, and thus it may be key to explain in future why HDAC6 deacetylates tubulins apart from histones as destined substrates of the HDAC family proteins. Remarkably, the very large majority of the high-confidence HDAC6 target sites map to IDRs, suggesting the possibility that IDR acetylation and its regulation by HDAC6 may generally be implicated in controlling LLPS. Notably, previous work reported that acetylated lysines are frequently located in regions with ordered secondary structure and, compared with all lysines, acetylated lysines are significantly enriched in structured regions and depleted in unstructured regions (Choudhary et al., 2009). Thus, the HDAC6 acetylome defines a skewed subset of the acetylated lysines, which may identify a new function for lysine acetylation. We focused on the HDAC6 target site in DDX3X among the high-confidence sites in IDRs. It is worth mentioning that, in addition to our study, an independent study also showed that DDX3X is acetylated

in mouse liver (Zhang et al., 2015b), indicating that the regulation of DDX3X by acetylation is observed not only in cell lines, but also in the whole organism.

One remaining question is what is the biological meaning of other high-confidence HDAC6 target sites? We realized that one of these is in the IDR of Serbp1, another SG component (Lee et al., 2014b), and thus the effect of acetylation/deacetylation was tested in a similar way as utilized for DDX3X. As expected, acetylation impaired localization of Serbp1 on SGs (**2.3.8 Apx. 8**, discussed in **3.2.5**), thus confirming the validity of our observations. Moreover, Ranbp2, which is part of the nuclear pore complex (nucleoporin) (Strambio-De-Castillia et al., 2010), may be also an attractive candidate for further analysis. As mentioned below, purified FG repeats of nucleoporins including Ranbp2 (also known as Nup358) were reported to form hydrogels through liquid-solid phase transition *in vitro* (Frey et al., 2006; Labokha et al., 2013), which is associated with sieve-like meshwork structure of the nuclear pore complex for selective shuttling of nuclear transport cargos in cells (discussed in **3.3.1**). Thus, in this case, HDAC6 may alter the material properties of the meshwork structure composed of Nups to control nuclear transport. Pex5 is the peroxisomal import receptor, which binds to its cargos in the cytosol and brings them to the translocation machinery at the peroxisomal membrane (Platta et al., 2007). Its deacetylation by HDAC6 was confirmed (**2.3.5 Apx. 5**). Synrg interacts with the AP1 clathrin-adaptor complex at the trans-Golgi network for trafficking of clathrin-coated vesicles (Page et al., 1999). Apart from the involvement of IDRs in phase separation, it has been also reported that IDRs contribute to trafficking; for example, IDPs drive membrane curvature to form clathrin-coated structures (Busch et al., 2015). Thus, Pex5 and Synrg, both proteins involved in the interaction between membranes and cargos may also play similar roles in the trafficking process with their IDRs, which would be modified by HDAC6. Clearly, further studies are required for each candidate identified here and another acetylome analysis with much higher resolution may be required in the future. However, the current list of novel HDAC6 substrate candidates presented here encourages us to imagine that HDAC6 globally deacetylates IDR, and controls the biological functions where IDRs play an important role, such as granule formation, establishment of sieve-like nuclear pore structure through

phase transitions and membrane curvature in the trafficking process.

3.2.2 DDX3X, as a component of stress granules and novel HDAC6 substrate

Our acetylome analysis revealed that DDX3X is a novel substrate of HDAC6 (see **Fig. 1 in 2.1**). DDX3X is a member of DEAD-box ATP-dependent RNA helicase family (Linder and Jankowsky, 2011). All DEAD-box helicases contain a structurally conserved helicase core with two RecA-like domains containing the characteristic Asp-Glu-Ala-Asp (DEAD) motif, whereas both N- and C-termini of the helicases are variable. DEAD-box helicases comprise the largest helicase family, with 37 members in humans and 26 in yeast. They are central players in essentially all aspects of eukaryotic RNA metabolism.

Ded1/DDX3 helicases form a defined subfamily, and the proteins are highly conserved from yeast to human (Sharma and Jankowsky, 2014; Soto-Rifo and Ohlmann, 2013). The human genome codes for two related DDX3 genes, DDX3X and DDX3Y. DDX3X was first identified as one of the X chromosome-linked genes (Park et al., 1998), and its functional homolog DDX3Y was found in the Y chromosome. DDX3Y is specifically expressed in testis and has an important function in spermatogenesis (Foresta et al., 2000). The helicase core harbors all signature helicase motifs (Fig. 16). Motif II (or Walker B motif) has the DEAD sequence, together with motif I (or Walker A motif), the Q-motif and VI, is required for ATP binding and hydrolysis. On the other hand, motifs Ia to Ic, III, IV, V and Va have been characterized less well but may be involved in interaction with RNA (Linder, 2006). In addition to this central helicase core, they have both characteristic N- and C-terminal domains, which are predicted as IDRs (Kim and Myong, 2016). These regions also provide potential sites for PTMs, most prominently the RG sequence in N-terminal IDR1 for methylation and the SR sequence in C-terminal IDR2 for phosphorylation (Gustafson and Wessel, 2010). However, there has been no report yet about DDX3X acetylation. While other DEAD box helicases, DDX5 and DDX17, are substrates of p300 (Mooney et al., 2010), no HATs were known for DDX3X acetylation.

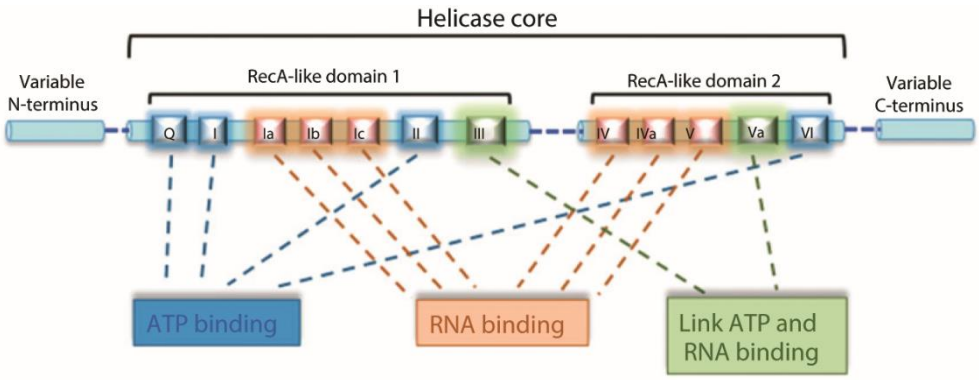


Figure 16. Functional domain organization of DEAD-box RNA helicase. Schematic representation showing domains and motifs common to the DEAD-box family. The catalytic helicase core is composed of two RecA-like domains, which contain conserved motifs involved in ATP binding (blue), RNA binding (brown), and linking of ATP and RNA binding (green). The variable N- and the C-terminus regions are also indicated. Adapted from (Soto-Rifo and Ohlmann, 2013).

Our results show that DDX3X-K118 is robustly acetylated by CBP and slightly by p300. DDX3X is ubiquitously expressed in multiple tissue types, and implicated in various processes of cellular RNA metabolism; it includes transcription regulation, pre-mRNA splicing, mRNA export, translation and cellular signaling. As expected from its versatile roles, DDX3X is also involved in human physiology such as viral infection, and its mutations are recurrently observed in some tumors. These characteristics of DDX3X are summarized schematically (Fig. 17) below, and discussed in more detail in the following section. Most importantly, LLPS propensity of DDX3X contributes to membrane-less organelle formation; this ability is the main focus of our study and discussed separately below (discussed in **3.2.5**).

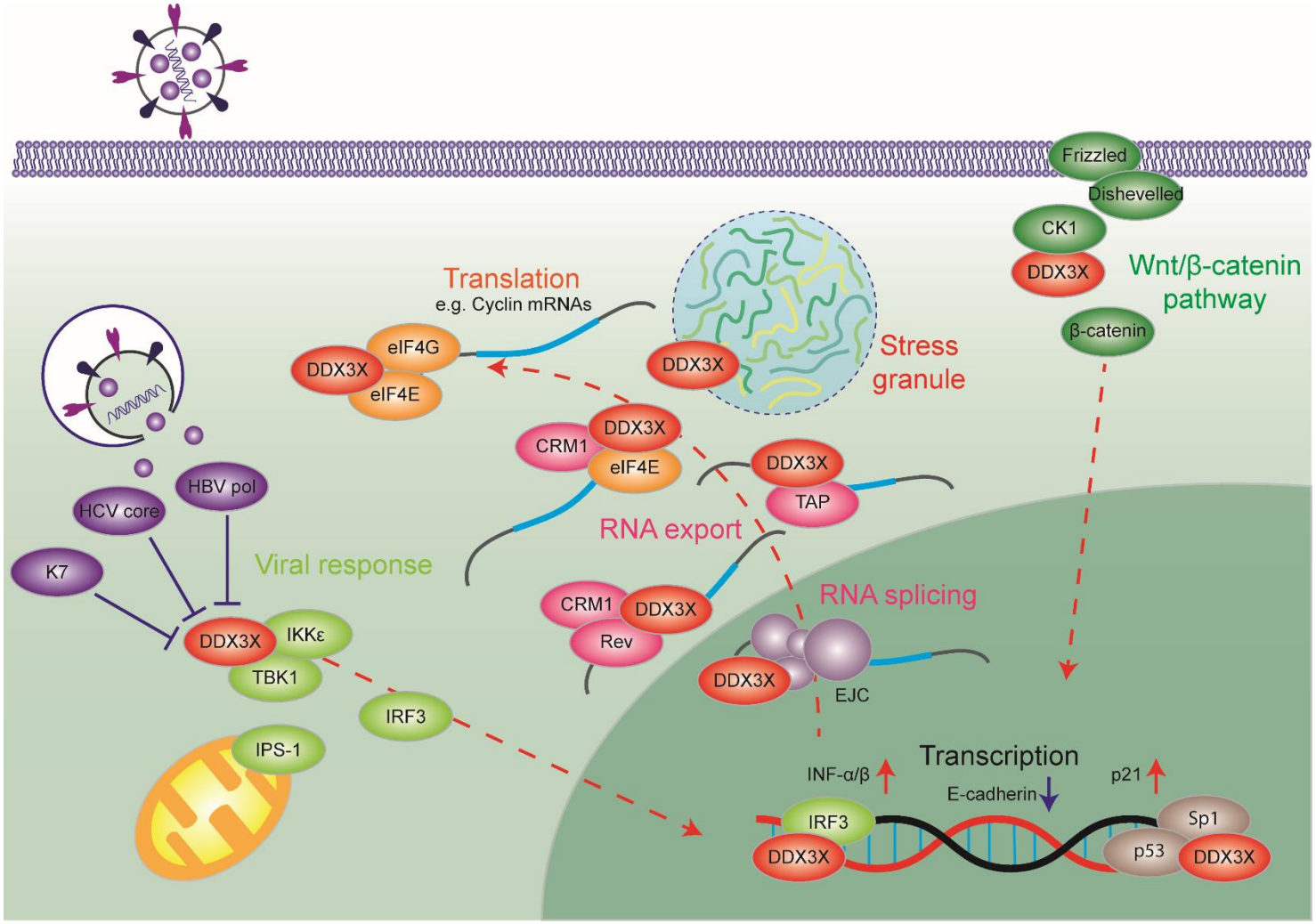


Figure 17. Diverse functions of DDX3X. The major biological processes regulated by DDX3X are presented: Transcription, RNA splicing, RNA export, Translation, Stress granule formation, Viral response and Wnt/β-catenin pathway.

RNA metabolism (transcription, processing and translation)

DDX3X is involved in transcription regulation for specific genes including p21, interferon-β (Inf-β), and E-cadherin, although it mainly localizes to the cytosol (Botlagunta et al., 2008; Chao et al., 2006; Soulat et al., 2008). DDX3X interacts with the transcription factor Sp1, and enhances transcription of p21 (Chao et al., 2006). Further studies directly indicated the association of DDX3X to the Inf-β promoter (Soulat et al., 2008), and DDX3X represses the expression from E-cadherin promoter in luciferase assay (Botlagunta et al., 2008). It still remains unknown how exactly DDX3X functions to regulate transcription of these selected genes.

In the process of RNA splicing, DDX3X associates with the exon junction complex (EJC) (Merz et al., 2007). The EJC binds to spliced mRNAs approximately 20 nucleotides upstream of splicing junctions in a sequence-independent manner, and plays important roles in mRNA quality control (Le Hir and Andersen, 2008). As the EJC is a highly dynamic, consisting of several core proteins and transient interactors, it is still unclear whether DDX3X functions with all or only selected EJCs. DDX3X is implicated in mRNPs (messenger ribonucleoproteins) export in three ways. First, DDX3X is implicated in CRM1-mediated export of intron-containing RNA from HIV (Yedavalli et al., 2004). The RNAs are bound by the HIV protein Rev, and DDX3X directly interacts with both the Rev and CRM1 to export RNA. Another study indicated that translation of the viral RNA is stimulated, suggesting the role of DDX3X in translation (Soto-Rifo et al., 2012). Second, DDX3X is involved in eIF4E/CRM1-dependent export mRNPs. A subset of mRNAs coding cell cycle regulators, such as cyclins, contain a characteristic RNA structure in their 3'UTR to efficiently bind to eIF4E. DDX3X associates with these mRNPs through protein-protein interactions (Topisirovic et al., 2009). Third, DDX3X also interacts with the mRNA export receptor TAP (NXF1) through its SR-rich IDR2 (Lai et al., 2008). DDX3X co-localizes with TAP in the nucleus and also in the cytosol. Thus, it is also implicated in TAP-dependent RNA export, which is distinct from the CRM1-mediated pathway.

DDX3X has been linked to some translation initiation factors, especially components of eIF4 (Shih et al., 2008; Soto-Rifo et al., 2012). DDX3X associates with eIF4E through consensus YxxxxLϕ

motif at its IDR1, and partially through IDR2. The direct interaction between eIF4G is also observed for yeast homolog Ded1p. As expected from its binding to translation initiation factors, DDX3X indeed regulates translation. Knockdown of DDX3X decreases production of β-globin from a reporter construct, without any alterations in mRNA processing events upstream of translation (Lee et al., 2008). DDX3X also promotes translation of a subset of mRNAs, notably the cyclin mRNAs, and viral RNAs with a stem-loop in their 5'UTR in whose export DDX3X is thought to be involved (Lai et al., 2008; Soto-Rifo et al., 2012). DDX3X is involved in cap-dependent translation (Geissler et al., 2012; Shih et al., 2008), and it promotes or inhibits translation depending on the context, as has been suggested for its yeast homolog Ded1p (Hilliker et al., 2011). A recent study proposed a model where Ded1p and its orthologues engage RNAs during the initiation of translation. Ded1p associates with the translation pre-initiation complex at the mRNA entry channel and affects translation initiation by controlling the use of near-cognate initiation codons that are proximal to mRNA structure in 5' untranslated regions (Guenther et al., 2018). A Ded1p mutant (with lower RNA binding ability, and impaired RNA unwinding ability) accumulates in 5' untranslated regions, resulting in the initiation of translation from near-cognate start codons there, and subsequent decrease in protein synthesis from the corresponding main open reading frames. This regulatory system may have a role in meiosis, where a marked decrease in the levels of Ded1p is accompanied by the activation of the alternative translation initiation sites that are observed with Ded1p mutant.

Cellular signaling

Apart from its main function as a regulator of RNA metabolism, DDX3X regulate cellular signaling, especially in two contexts: induction of IFN-α/β upon virus infection and the Wnt/β-catenin signaling pathway.

The IKKε/TBK1 complex is activated by upstream signaling cascade originating from pattern recognition receptors such as Toll-like receptors upon viral infection and leads activation of IRF3 and production of IFN-α/β (Kawasaki and Kawai, 2014). DDX3X interacts with IKKε and promotes its

autophosphorylation and activation (Gu et al., 2013). IKKε phosphorylates several serine residues in DDX3X IDR1, and these phosphorylation promotes IRF3 to bind DDX3X, suggesting that DDX3X functions as adaptor that enhances activation of IRF3 by IKKε. Another study demonstrated that DDX3X also interacts with TBK1 and that its helicase core is phosphorylated (Soulat et al., 2008). This phosphorylation is important for DDX3X on the IFN-β promoter to mediate its transcription, although the involvement of IRF3 in this process was not tested. DDX3X also regulates the mitochondrial protein IPS-1, and functions as a regulator of IKKε/TBK1 complex, in order to activate IRF3 (Oshiumi et al., 2010).

In the Wnt/β-catenin signaling pathway, DDX3X acts as regulatory subunit of casein kinase 1 (CK1) (Cruciat et al., 2013). CK1 phosphorylates and activates the scaffold protein dishevelled (Dvl), which binds components of Wnt receptor and the β-catenin destruction complex (Wallingford and Habas, 2005). Upon activation of Wnt receptors, DDX3X binds to and allosterically activates CK1. The helicase activities of DDX3X, including ATP hydrolysis and RNA unwinding, are not required for its ability to bind CK1, but its C-terminal region including IDR2 is required. DDX3X knockdown impairs the developmental process also in *X. tropicalis* and *C. elegans*, suggesting that the function of DDX3X in Wnt/β -catenin signaling is evolutionary conserved.

Roles in human physiology

DDX3X expression is deregulated in various types of tumors, and its recurrent mutations are associated with some specific types of tumors. Cancer promoting and suppressing roles of DDX3X are dependent on tumor cell types. DDX3X mRNA is overexpressed in hepatocellular carcinoma (HCC) tissues and, overexpression of DDX3X induces anchorage-independent growth of cells (Huang et al., 2004). However, another study demonstrated downregulation of DDX3X in HCC tissues (Chang et al., 2006a). Although the mechanism is unclear, DDX3X supports the interaction between p53 and Sp1 to promote the expression of p53-target genes such as p21, which is deregulated in lung tumors (Wu et al., 2011). The most persuasive evidence that DDX3X is important for tumor initiation and/or

progression is coming from recent whole-exome sequencing studies of different cancer samples. Mutations are recurrently observed at catalytically important amino acids located in the DDX3X helicase core in a spectrum of cancers: chronic lymphocytic leukemia (Wang et al., 2011), head and neck squamous cell carcinoma (Stransky et al., 2011), medulloblastoma (Pugh et al., 2012; Robinson et al., 2012), and natural killer/T-cell lymphoma (Jiang et al., 2015). In some case of medulloblastoma, DDX3X and β-catenin mutations simultaneously occur. Expression of DDX3X mutant and β-catenin mutant indeed enhances the Wnt signaling activity defined by a reporter assay (Pugh et al., 2012). Given that DDX3X regulates β-catenin through CK1 activation, it may be possible that cancer-related mutations affect their interaction. The exact molecular mechanisms by which DDX3X mutants activate Wnt/β-catenin signaling pathway remain to be clarified in future, and moreover it should also be revealed whether these recurrent DDX3X mutations are driver mutations for the oncogenesis, or just support tumor progression.

Diverse viruses utilize DDX3X as cellular co-factor for their replication. DDX3X binds to HIV Rev proteins, and is involved in export and translation of HIV RNAs, as mentioned above. And the helicase activity of DDX3X is required for HIV replication (Yedavalli et al., 2004). This was further confirmed that knockdown of DDX3X inhibits the export of HIV RNAs and viral replication without affecting cell viability (Ishaq et al., 2008), suggesting the possibility of DDX3X as a drug target. DDX3X is also involved in the replication of Hepatitis C virus (HCV); two genome-wide screens identified DDX3X as a cellular co-factor of HCV replication. DDX3X interacts and co-localizes with the Hepatitis C Virus (HCV) core protein, which forms part of the viral nucleocapsid, and it is required for HCV replication (Ariumi et al., 2007; Owsianka and Patel, 1999). It is still unknown how the interaction between HCV core protein and DDX3X supports viral replication. One possibility is that sequestering of DDX3X by the HCV core protein might disrupt other cellular functions of DDX3X, such as IFN-β production as mentioned above. This was observed in the case of another hepatotropic virus, Hepatitis B Virus (HBV); HBV polymerase binds to DDX3X, and blocks the function of DDX3X in innate immune signaling pathways (Wang and Ryu, 2010). In the case of vaccinia virus

(VACV), the viral protein K7 binds to DDX3X, and inhibits the same pathways (Schroder et al., 2008). DDX3X levels were found to be reduced in HBV-positive HCC samples, while no reduction was observed in HCV-positive samples (Chang et al., 2006a), indicating that manipulation of DDX3X by these two viruses is different. Recent studies demonstrated that DDX3X also positively regulates Norovirus and West Nile virus replication (Chahar et al., 2013; Vashist et al., 2012), indicating that interactions between DDX3X and viral proteins for replication is a more general feature of RNA viruses.

DDX3X is also related to a kind of intellectual disability (ID), which is called as "DDX3X syndrome" (Snijders Blok et al., 2015). Whole exome sequencing revealed that mutations in DDX3X helicase core are observed in female ID patients who also show various other features including hypotonia, movement disorders, behavior problems, corpus callosum hypoplasia, and epilepsy. These DDX3X mutation may account for 1-3% of unexplained ID in females. As mentioned above, DDX3X activates Wnt/β-catenin signaling pathway through the interaction with CK1. It was known that exogenous Wnt expression results in ventralization of zebrafish embryos, and this effect is enhanced by overactivation of Wnt due to DDX3X co-expression. The ID-related mutants were tested in this assay, and indeed exhibit a consistent loss-of-function effect on the Wnt/β-catenin pathway, suggesting that the mutations in DDX3X helicase core don't efficiently activate the pathway, and that impaired Wnt/β-catenin pathway may cause ID.

3.2.3 Acetylation in stress response

We discovered that CBP autoacetylation occurs immediately after cells are treated with stress inducers (for oxidative stress, puromycin and sorbitol), resulting in global protein acetylation including DDX3X (see **Fig. 2 in 2.1**). A few studies have suggested a role for lysine acetylation on granule formation under stress, although without mechanistic insight. An RNAi screen in human cells identified the HAT p300 as an important factor for SG formation, although the outcome whether SG formation is inhibited or enhanced was not mentioned clearly (Ohn et al., 2008). Another study

described the deacetylase SIRT6 as a critical factor for the process in *C. elegans*, and also in mammalian cells; SIRT6 knockout MEFs form less large SGs compared to WT under heat shock condition (Jedrusik-Bode et al., 2013). They demonstrated that the deacetylase activity of SIRT6 regulates G3BP dephosphorylation at S149 by unknown mechanism. As G3BP phosphorylated at S149 shows impaired self-association activity (Tourriere et al., 2003), they proposed that this would lead a defect in SG formation. Their result – knockout of a specific HDAC disrupts large stress granule formation – is fully compatible with our observations here and in the past (Kwon et al., 2007). It may be possible that knockout of SIRT6 also increases the acetylation in IDR-containing proteins, and controls their LLPS for SG formation besides the effect on G3BP. A recent study in *S. cerevisiae* demonstrated that the KAT complex NuA4 enhances SG formation upon glucose deprivation but not heat stress, and furthermore that pharmacological inhibition of Tip60, the human homolog of NuA4, decreases SGs in human cell lines (Rollins et al., 2017). The authors did not mention any effector molecules acetylated by NuA4/Tip60, but instead they showed that NuA4 is important for the full-activity of Ac-CoA carboxylase Acc1. As Acc1 converts Ac-CoA to malonyl-CoA, NuA4 inhibition induces the accumulation of Ac-CoA in cells. They proposed that elevated Ac-CoA levels suppress SG formation, whereas decreased Ac-CoA levels enhance it upon glucose deprivation by an unknown mechanism. At first glance, the positive effect of a KAT NuA4 on SG formation seems to contradict the model that acetylation inhibits LLPS/SG formation. However, their work showed that Ac-CoA level increases under NuA4 inhibition, suggesting that increased Ac-CoA in this context may further be used for acetylation of IDR-containing proteins by other KATs. Although the studies mentioned above claimed the importance of catalytic activities of KATs/KDACs for SG formation, molecular mechanisms including their substrates were not described. A study showed that doxorubicin-mediated oxidative stress increased p300 expression in cardiac myocytes through its autoacetylation (Jain et al., 2012). The increase in acetylated p300 is likely due to p300 autoacetylation, as it is decreased by inhibitors of p300 KAT activity, which is consistent with the observation that acetylated p300 enhances resistance to proteasomal degradation. In parallel, p300 mRNA decreased significantly over the time-

course, indicating that transcriptional induction is unlikely to account for the robust increase in p300 during stress. This finding raises an important physiological mechanism to control p300 activity in the response to acute oxidative stress, and is a good match with our observations about CBP. Acetylation regulates RNP granule formation not only through altering net charge of IDR, but also through changing other protein properties. For example, acetylation in the RNA binding domain of TDP-43 impairs its RNA binding ability and enhances aggregate formation, showing another critical role of acetylation in RNP granule formation (Cohen et al., 2015). In this case, acetylation sites are not in the IDR of TDP-43, but in its RNA binding domain and may affect the protein stability of TDP-43 or its association with RNA (discussed in **1.2.2**). There are multiple open questions in this physiological system to control RNP granule formation through stress-regulated acetylation. Our study revealed that DDX3X acetylation by CBP is increased in early phase after stress-inducer treatment, while deacetylation by HDAC6 in SGs occurs in later phase. Along the same line, the kinetics of each KAT/KDAC under several types of stress should be examined in the future, and how specific stresses lead to their activation with concomitant acetylation of multiple targets should also be investigated. Moreover, it still remains unknown if/how the granule formation ability of other RNA binding proteins and RNA helicases are regulated by acetylation. Most importantly, it is currently unknown whether or not this global acetylation provides any benefits to cellular fitness. Nevertheless, lysine acetylation and deacetylation, as a regulator of RNP granule formation, may offer a novel therapeutic rationale for targeting HDACs or HATs in diverse neurodegenerative diseases.

3.2.4 Roles of posttranslational modifications in phase separation

Membrane-less organelle formation can also be regulated by modulating the phase separation threshold by changing the degree of valency and intrinsic solubility of components; PTMs appear to be an important factor to control these properties. We found that DDX3X-IDR1 can undergo LLPS, and lysine acetylation in the IDR1 disrupts it (see **Fig. 3 in 2.1**). Positively charged lysine residues in an IDR contribute to cation-anion or cation-π interactions, which are driving forces of LLPS. Lysine acetylation weakens these interactions due to neutralization, and as a result, LLPS of the IDR is attenuated. Our data also demonstrate that spatiotemporal modifications of DDX3X-IDR1 are regulated by CBP and HDAC6 *in vivo*, and that (de)acetylation-modulated DDX3X LLPS contributes to SG formation. As far as we know, DDX3X is the first example where acetylation regulates LLPS; (de)acetylation-modulated LLPS may become a general mechanism to control the formation of membrane-less organelles. Several examples of the phase separation controlled by PTMs have been recently revealed; they comprise phosphorylation, methylation, sumoylation, non-covalent interactions between ubiquitin/polyubiquitin and their receptor proteins, and Poly(ADP-Ribose), as well as acetylation first revealed by our study (Table 5).

PTMs	Targets	Molecular Determinants	Effect on phase separation	Reference
Phosphorylation	N-WASP signaling	MD	Enhanced	(Li et al., 2012)
Phosphorylation	LAT signaling	MD	Enhanced	(Su et al., 2016)
Phosphorylation	HP1α	IDR	Enhanced	(Larson et al., 2017)
Phosphorylation	PGL-1	N.A.	Enhanced	(Zhang et al., 2018a)
Phosphorylation	Synapsin1	IDR/MD	Attenuated	(Milovanovic et al., 2018)
Phosphorylation	FUS	LCS	Attenuated	(Monahan et al., 2017)
Phosphorylation	TDP-43	α-helix	Attenuated	(Wang et al., 2018a)
Phosphorylation	DYRK substrates	N.A.	Attenuated	(Rai et al., 2018)
Phosphorylation	RNA Pol II CTD	LCS	Enhanced/ Attenuated	(Kwon et al., 2013; Lu et al., 2018)
Methylation	DDX4	RGG/RG motif	Attenuated	(Nott et al., 2015)
Methylation	hnRNPA2	RGG/RG motif	Attenuated	(Ryan et al., 2018)
Methylation	FUS	RGG/RG motif	Attenuated	(Hofweber et al., 2018; Qamar et al., 2018)
Methylation	PGL-1, PGL-3	RGG/RG motif	Attenuated	(Zhang et al., 2018a)
Sumoylation	PML body components	MD	Enhanced	(Banani et al., 2016)
Ubiquitin chain	UBQLN2	IDR	Attenuated	(Dao et al., 2018)
Ubiquitin chain	p62/SQSTM1	MD	Enhanced	(Sun et al., 2018)
Poly(ADP-ribose)	TDP-43	N.A.	Enhanced	(McGurk et al., 2018)
Acetylation	DDX3X	LCS	Attenuated	(Saito et al., 2019)

Table 5. Summary of phase separation regulated by PTMs. Molecular determinants indicate expected driving force for phase separation, that is affected by PTMs. MD: modular interaction domains; N.A.: not assessed.

Phosphorylation

Phosphorylation functions as ligand for binding modules of multivalent proteins, and thus promotes LLPS. Actin-regulatory signaling pathway, composed of nephrin, Nck and neural Wiskott-Aldrich syndrome protein (N-WASP), is regarded as the first example that phosphorylation plays an important role in LLPS (Li et al., 2012) (discussed in **1.3.2**); these components form oligomers through

SH2 domains and phosphotyrosine residues (ligands of SH2 domains). LAT signaling assembly is also controlled mainly by interaction between SH2-containing proteins (GRB2, GADS and SLP76) and phosphortyrosine residues on the components (LAT and SLP76) (Su et al., 2016) (discussed in **1.3.2**). Phosphorylation also alters the electrostatic and/or oligomerization abilities of IDRs, and affects phase separation. The N-terminal IDR of HP1α is phosphorylated, which induces its conformational change and subsequent higher-order oligomerization, and results in enhanced LLPS (Larson et al., 2017). Similarly, mTORC1-mediated phosphorylation of PGL-1 and PGL-3 is important for these proteins to phase separate and form PGL granules, which confer heat-stress resistance on *C. elegans* embryos, although the molecular mechanisms of their LLPS remain to be solved (Zhang et al., 2018a). It was also shown that phosphorylation can also negatively regulate phase separation. The C-terminal IDR of synapsin1 is phosphorylated by CaMKII, and this phosphorylation disrupts LLPS of synapsin1 to release lipid vesicles, mimicking the process observed at presynaptic sites (Milovanovic et al., 2018). Phase separation of neurodegenerative diseases-associated proteins is also regulated by phosphorylation, showing therapeutic possibilities to target kinases and phosphatases for the treatment. The N-terminal LCS of FUS is enriched with serine and threonine residues, which are phosphorylated by DNA-PK, and resulting negative charges at S/T sites in the LCS inhibit phase separation (Monahan et al., 2017). TDP-43 has a single serine phosphorylation site in the N-terminal globular domain for its oligomerization. Phosphorylation impairs the oligomerization and LLPS propensities, and increases fluidity of its nuclear assembly (Wang et al., 2018a). Although these examples clearly demonstrate the importance of phosphorylation in phase separation, a limited number of studies have focused on the interactome of a certain kinase or phosphatase in the context of membrane-less organelle formation. An exemplary study explained the molecular mechanism that membrane-less organelles disappear upon nuclear-envelope breakdown and reappear when mitosis is completed, and demonstrated that dual-specificity kinase DYRK3 acts as a central “dissolvase” for this process (Rai et al., 2018). The DYRK3 interactome was obtained by SILAC, which revealed that the majority of proteins identified were known to bind RNA, and also include well-known components of biomolecular condensates such

as SGs, splicing speckles and the centrosome or pericentriolar matrix. Pharmacological inhibition of DYRK3 kinase activity causes aberrant mitotic hybrid phase-separated components, indicating that DYRK3 is important to prevent the unmixing of the mitotic cytoplasm. The other well-known example to show phosphorylation-mediated phase separation is a subunit of RNA polymerase II (Pol II) (discussed in **3.3.2**).

Methylation

Positively charged arginine residues in an IDR contribute to cation-anion or cation-π interactions for phase separation. Unlike lysine acetylation, arginine methylation does not affect the positive net charge, but it alters hydrogen bonding and local hydrophobicity of arginines (Fuhrmann et al., 2015), and thus influences cation-π interactions that drive phase separation. The N-terminal IDR of DDX4 alone is sufficient for LLPS, and it contains several RGG/RG motifs where arginine methylation possibly occurs (Nott et al., 2015). Arginine methylation is regulated by protein arginine methyltransferases (PRMTs) *in vivo*. Several independent proteomic analyses identified arginines in LCS of hnRNPA2 that are methylated by PRMT1; two studies found demethylation and monomethylation at several modified RGG sites (Friend et al., 2013; Geoghegan et al., 2015), while another study further suggested that PRMT1 can methylate arginines at non-RGG sites, including RGF, which is also observed in the hnRNPA2 LCS (Wooderchak et al., 2008). Methylation of hnRNPA2 LCS by PRMT1 *in vitro* induces asymmetric dimethylation of the several RGG sites there, and indeed decreases LLPS (Ryan et al., 2018). This RGG/RG motif is often found in RNA-binding proteins; another ALS-related protein, FUS, also has RGG/RG-rich domains at its C-terminus (CTD) in addition to the N-terminal LCS. Arginine methylation of RGG/RG motifs reduces phase separation of FUS and increases droplet dynamics. Moreover, hypomethylation of FUS caused by adenosine-2,3-dialdehyde (AdOx), an inhibitor of PRMT, is accompanied with enhanced cellular FUS granules formation in an AdOx dose-dependent manner (Qamar et al., 2018). It has also been revealed that TNPO1 (also known as importinβ2) functions as a chaperone protein for FUS (Guo et al., 2018a; Yoshizawa et al., 2018).

The chaperone activity of TNPO1 is not impaired when FUS is hypomethylated, and TNPO1 has an even higher affinity for hypomethylated FUS (Hofweber et al., 2018), suggesting that TNPO1 efficiently prevents FUS from phase separation when the methylation level of FUS is decreased. As mentioned above (discussed in **1.3.2**), repetitive -G/S-Y- motifs in FUS N-terminal LCS are prominent drivers of its LLPS. FUS phase separation is impeded by either replacement of tyrosines in N-terminal LCS with alanines or arginines in CTD by alanines (Qamar et al., 2018). This strongly suggests that aromatic π of tyrosine residues in the LCS of FUS interact with cationic arginines in its CTD inter- and intra-molecularly, and thus that these cation-π pairings contribute to phase separation of FUS. This regulatory mechanism of hnRNPA2 and FUS by methylation will perhaps encourage the development of drugs that specifically activate PRMTs in ALS patients. Furthermore, methylation-mediated LLPS plays a physiological role in *C. elegans* embryos apart from pathological context; methylation of PGL-1 and PGL-3 attenuates LLPS of PGL granules (Zhang et al., 2018a). This methylation simultaneously enhances their selective autophagic degradation to restrict germline P granules to germ blastomeres during embryogenesis (Li et al., 2013; Zhang et al., 2009).

Ubiquitin-like modification

Promyelocytic leukemia (PML) body is a membrane-less organelle that has been implicated in the regulation of various cellular functions, such as the induction of apoptosis and cellular senescence, inhibition of proliferation, maintenance of genomic stability and antiviral responses (Bernardi and Pandolfi, 2007). A large part of PML bodies components, including its prominent component PML, are sumoylated, and SUMO-SIM interactions among these components are considered as a driving force for the formation of PML bodies. Although a reconstituted system composed of synthetic SUMO and SIM proteins provided the evidence that valency of these two components defines the composition of PML body *in vitro* (Banani et al., 2016), the role and dynamics of each PML body component *in vivo* still remain unknown. It is also unknown whether or not ubiquitination affects LLPS, but there is some evidence showing that ubiquitin or polyubiquitin chains

affect LLPS. Proteasomal shuttle factor UBQLN2, which mediates autophagy-independent protein aggregate clearance by the proteasome, is a component of SGs. UBQLN2 has an IDR at its C-terminus including oligomerization domain, and it exhibits LLPS *in vitro* (Dao et al., 2018). The oligomerization ability is connected to LLPS ability and its localization in SGs. Ubiquitin or polyubiquitin binding to UBA domain of UBQLN2 attenuates its LLPS, thus serving as a switch between condensed and dispersed phases. Hence, UBQLN2 may be recruited to SGs through its intrinsic LLPS ability without any ubiquitin binding, where its interactions with ubiquitinated substrates may reverse their LLPS to enable shuttling of clients out of SGs. LLPS mediated by ubiquitin is also important in protein aggregate clearance by autophagy. An autophagy receptor p62/SQSTM1 interacts with both ubiquitin and LC3/GABARAP on the autophagosome, thereby promoting autophagy of ubiquitinated aggregated proteins (Kirkin et al., 2009) (discussed in **1.2.3**). The current concept is that ubiquitin-tagged misfolded proteins are assembled into aggregates by the scaffold protein p62, and these aggregates are then engulfed and degraded by autophagosomes. One study revealed that p62 forms liquid-like droplets *in vivo* (Herhaus and Dikic, 2018; Sun et al., 2018). Purified p62 itself doesn't undergo LLPS *in vitro*, but adding a K63 polyubiquitin chain to p62 induces its LLPS, indicating that the p62 LLPS is mechanistically dependent on its polymerization through the interaction between p62 and ubiquitin/polyubiquitin. p62 is mutated in Paget's disease, and disease-related mutations impaired its LLPS. These examples clearly demonstrate that the LLPS propensities of proteasome shuttle factors or autophagy receptors are mediated by ubiquitin/polyubiquitin binding, and that their LLPS plays an important role for aggregate clearance.

Poly(ADP-ribose)

Poly(ADP-ribose) (PAR) is a negatively charged biopolymer, which is covalently attached to target proteins by PAR polymerases (PARPs); it is recognized by specific ‘‘reader’’ proteins, and thus, PAR drives the assembly of protein complexes. SG and aggregate formations of TDP-43 are regulated by PAR; TDP-43 binds to PAR through its nuclear localization sequence, and this binding promotes

its LLPS (McGurk et al., 2018). As PAR localizes in SGs, it functions as a homing signal for TDP-43, directing it to SGs. In Drosophila models of neurodegenerative diseases, reduction of a representative PARP called tankyrase, mitigates the disease phenotype associated with TDP-43 overexpression. Consistently, tankyrase inhibitors, which have been developed as cancer therapeutics, can reduce SG accumulation of TDP-43 in mammalian cells, likely by inhibiting PARylation of SG components and allowing TDP-43 to shuttle back to the nucleus; this suggests that tankyrase inhibition could be developed as a therapeutic modality for TDP-43 associated ALS/FTD.

Acetylation

Acetylation can be now added to the list of PTMs that regulate phase separation and membrane-less organelle formation on the basis of our work (see **2.1**). We observed that DDX3X-IDR1 can undergo LLPS, and that lysine to glutamine (KQ and allQ) mutants exhibited reduced LLPS propensity. Positively charged lysine residues in an IDR contribute to cation-anion or cation-π interactions, which are driving forces of LLPS. Lysine acetylation weakens these interactions due to neutralization, and as a result, LLPS of the IDR may be attenuated. We also examined systematic K to Q mutagenesis at multiple lysines in DDX3X-IDR1 to test its LLPS propensity, and found interesting positional effects of lysines there. Mutation on seven residues (K35, 50, 55, 64, 66, 81, and 118) impaired LLPS, while three residues (K130, 138 and 162) did not (Fig. 18). One study redefining the DDX3X catalytic core may explain these positional effects of K to Q mutations on LLPS (Floor et al., 2016). The DDX3X subfamily has a conserved N-terminal extension (NTE, a.a. 132-167 in human and mouse), which harbors a predicted short α-helix at the C-terminal end of the region (a.a. 1-167) that is predicted to be an IDR by several established algorithms. Although secondary structures such as α-helix were also reported to trigger LLPS (Conicella et al., 2016), our observation indicates that in the case of DDX3X lysine acetylations in a “genuine” IDR (K35, 50, 55, 64, 66, 81, and 118) impair LLPS more effectively than acetylations in an α-helix. We also confirmed that cortactin, a well-established substrate of HDAC6, also has an IDR in its N-terminus (Li et al., 2017), and undergoes

LLPS. In this case as well, K to Q mutation, as an acetylation mimic, impairs LLPS. Thus, our observations with DDX3X hint to a general mechanism for LLPS regulation by acetylation. In addition to impaired droplet formation, we observed that acetylated DDX3X-IDR1 is not precipitated from cellular lysates by the b-isox chemical. It would be interesting to determine the structure of soluble, monomeric DDX3X-IDR1 by solution NMR spectroscopy (as done in (Monahan et al., 2017; Ryan et al., 2018)) and compare with the acetylated form.

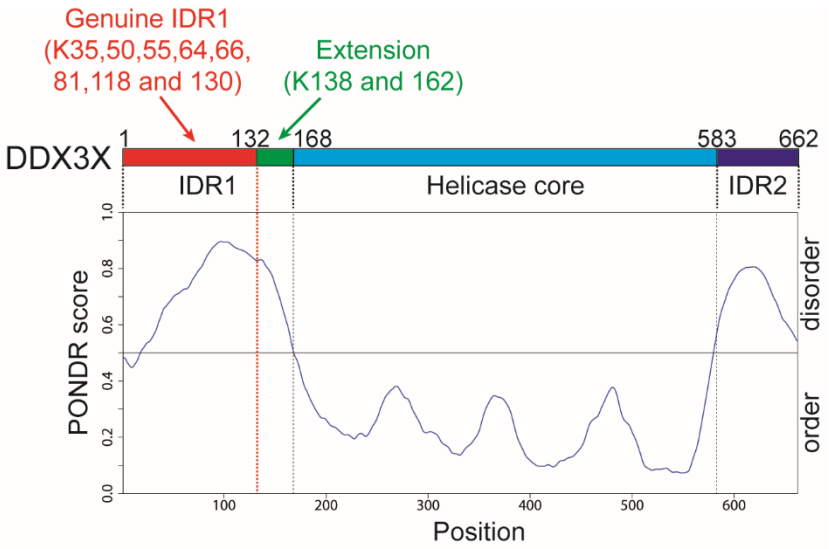


Figure 18. Positional effect of acetylation on DDX3X-IDR1. Although the structure prediction program VL3-BA defines DDX3X a.a. 1-167 as an N-terminal IDR1 (bottom, PONDR score graph), a recent structural analysis study (Floor et al., 2016) defines only a.a. 1-131 as an IDR1 of DDX3X. The region (a.a.1-131) in red represents this genuine IDR1, while the region in green (a.a. 132-167) represents an α-helical extension of the helicase core. All 10 lysine residues in each region are numerated.

3.2.5 DDX3X in stress granule formation

DDX3X has been reported as a component of SGs and neuronal kinesin-associated granules (Kanai et al., 2004; Lai et al., 2008; Shih et al., 2012). Overexpressed DDX3X-GFP formed SGs (Lai et al., 2010), which reflects the SG formation observed in yeast upon overexpression of DDX3X homolog Ded1p (Beckham et al., 2008). This study reported that knockdown of DDX3 did not impair

SG assembly under arsenite (Lai et al., 2010), while the other reported that knockdown of DDX3X interfered with SG formation under the same context (Shih et al., 2012). The difference between these two conditions may be due to knockdown efficiency of DDX3X and/or SG-quantification methodology. We established DDX3X KO MEFs and observed that the total SG area defined by G3BP immunofluorescence microscopy in cell is decreased in KO cells under specific stress conditions (oxidative stress, energy depletion and puromycin). As DDX3X can be divided into three functional domains, IDR1, helicase core and IDR2, it is important to dissect the contribution of each domain to SG formation. While the role of IDR1 was extensively studied in our work (**2.1**), several observations argue that RNA/DNA helicases including DDX3X, which utilize the energy of ATP hydrolysis to unwind DNA/RNA, play roles in controlling SG formation and disassembly. This supports a concept of SG as 'active liquids' where the energy of ATP-driven remodeling keeps the assembly in a dynamic state. Indeed, previous work has already tried to dissect the role of the helicase core for its SG localization and/or formation, but it is still unresolved; mutations in important residues for its helicase activity in DDX3X do not interfere with SG formation (Shih et al., 2012), while medulloblastoma-associated mutations in its helicase core enhance SG formation (Valentin-Vega et al., 2016). These studies only provide qualitative data, that is, percentage of SG-positive population, and do not quantify SG parameters such as its number and average size inside each cell. Thus, we also tested the contribution of DDX3X helicase core domain to SG formation with our experimental setup with quantification (**2.3.6 Apx. 6**). DQAD (E348Q) mutant and AAA (S382A and T384A) mutant are commonly used in the field; a point mutation of the DEAD sequence (Motif II) to a DQAD severely impairs ATPase activity and helicase activity, while the conversion of the SAT motif (Motif III) into AAA impairs the RNA-unwinding activity, but retains the ATPase activity of the helicase (Soto-Rifo and Ohlmann, 2013). A study about the helicase core of another DEAD-box helicase, Vasa (DDX4 ortholog of *D. melanogaster*), has revealed other key residues for its function (Sengoku et al., 2006). As DEAD-box helicase cores are well conserved between DDX3X and Vasa, the corresponding four residues of DDX3X: Q477A and R480A (Motif IVa), T498A (Motif V) and R503A (Motif Va) are

mutated, and their SG formation ability was also tested (**2.3.6 Apx. 6**). All six mutants formed smaller SGs, confirming that the helicase core is important for normal SG formation. Increased SG number was also observed with DQAD and AAA mutants, and DQAD mutant showed the most significant difference compared to other constructs. Although four mutants (Q477A, R480A, T498A and R503A) lose RNA-binding ability (**2.3.6 Apx. 6C**), DQAD and AAA mutants are still fully capable to bind RNA (Pause and Sonenberg, 1992). Thus, the RNA-binding ability itself may not be the cause of impaired SG formation observed in this experiment. Given that all mutants exhibit impaired RNA-unwinding ability, this primary helicase function seems to be connected to maintenance of SG rather than simple SG disassembly, and may give a good explanation for the observation here. Although both DQAD and AAA mutants lost RNA-unwinding ability, ATPase activity-deficient DQAD mutant formed much smaller SGs than ATPase activity-intact AAA mutant, implying that ATPase activity itself contributes to keep SG size. It is important to mention that DDX3X plays important roles in regulation of translatome as mentioned above, and other RNA-binding proteins, ribosomes and RNAs may be affected by DDX3X mutants, resulting in impaired SG formation. Obviously, further systematic studies to compare in parallel several biological function of DDX3X including three major properties of the helicase core (RNA-binding, ATPase and RNA-unwinding), global effect on translatome and SG formation would be required.

On the other hand, the contribution of N-terminal IDR1 has been firmly confirmed independently by several groups (Shih et al., 2012; Valentin-Vega et al., 2016). Moreover LAF-1, a related *C. elegans* protein, has been shown to undergo LLPS, which depends on its N-terminal RGG region (Elbaum-Garfinkle et al., 2015). Here, we demonstrated that DDX3X-IDR1 undergoes LLPS, and that this propensity is important for both localization of DDX3X into SGs and their maturation. In addition to acetyl-mimic/dead mutations, cancer-related mutations in IDR1 also affected its LLPS (**2.3.7 Apx. 7**). Especially arginine mutations in RG or RXG sequences impaired its LLPS, suggesting that these motifs function as driving force for LLPS also in the case of DDX3X-IDR1. The biological relevance of these mutations and stress sensitivity in cancer cells should be tested in near future.

Spatiotemporal modifications of DDX3X-IDR1 regulated by CBP and HDAC6 control SG maturation. Earlier work from our laboratory (Kwon et al., 2007) demonstrated that loss of HDAC6 reduces the number of large SGs under arsenite treatment, which is likely related to the work presented here. It is important to note that other SG components harboring IDRs also would possibly be regulated by HDAC6. We realized that one of our HDAC6 high-confidence target sites is in IDR of Serbp1, another SG component (Lee et al., 2014b). Serbp1, a RNA binding protein, harbors a long IDR containing two RGG/RG-rich sequences (RGG box1 and 2). These RGG/RG-rich regions are also methylated by PRMT1 as mentioned above (discussed in **3.2.4**), and hypomethylated Serbp1 prefers to localize in the nucleus (Lee et al., 2012). Along the same line, we observed that large part of Serbp1 SG formation ability is depending on these two RGG boxes, suggesting the importance of these regions for its SG formation (**2.3.8 Apx. 8B**). As observed in DDX3X, acetyl-mimic mutation (K122Q) in the Serbp1 IDR also affected SG formation; numerous smaller SGs were induced (**2.3.8 Apx. 8C**). Thus, HDAC6 may regulate the acetylation status of IDRs in a set of proteins contributing to SG formation.

Several models have been proposed to explain the formation of SGs. A prevalent "Cores First" model (Jain et al., 2016; Protter and Parker, 2016; Wheeler et al., 2016) suggests that mRNPs first condense into stable core structures through strong specific interactions (initiation), and the high local concentration of IDRs on SG components then triggers a LLPS for shell components of SGs (maturation). Specific stresses induce autoacetylation and activation of CBP, and lead to global protein acetylation, including DDX3X. Another RNP component, TDP-43, also becomes acetylated by CBP under oxidative stress, and aggregated. This aggregation is resolved by counteraction of HDAC6 (discussed in **1.2.2**). Many open questions in this granules/aggregates formation mediated by stress-induced acetylation still remain unanswered (discussed in **3.2.3**). Nevertheless, active deacetylation by HDAC6 of some RNP components is essential for their normal dynamics and function in the stress response; deacetylation is necessary for some IDR-containing proteins including DDX3X to undergo LLPS and induce appropriate SGs formation. Our quantitative analysis and mathematical modeling of SG growth with a series of DDX3X mutants (see **Fig. 5 in 2.1**) provided additional evidence that

regulation of LLPS by acetylation/deacetylation is critical not for the initiation of the SG formation, but for their maturation, which is consistent with the model mentioned above. The comparison of interactomes between DDX3X WT and acetyl-mimic allQ mutant revealed that the acetyl-mimic mutant does not interact with some IDR-containing proteins such as hnRNPA2/B1. This further shows that the recovery of DDX3X LLPS propensity by HDAC6 on SGs is important for localized recruitment of other proteins which themselves can undergo LLPS and contribute to the formation of SG shell together with DDX3X. Thus, our results demonstrate a novel physiological role for acetylation/deacetylation in LLPS during SGs maturation (Fig. 19), and may help to further refine the models for SG assembly.

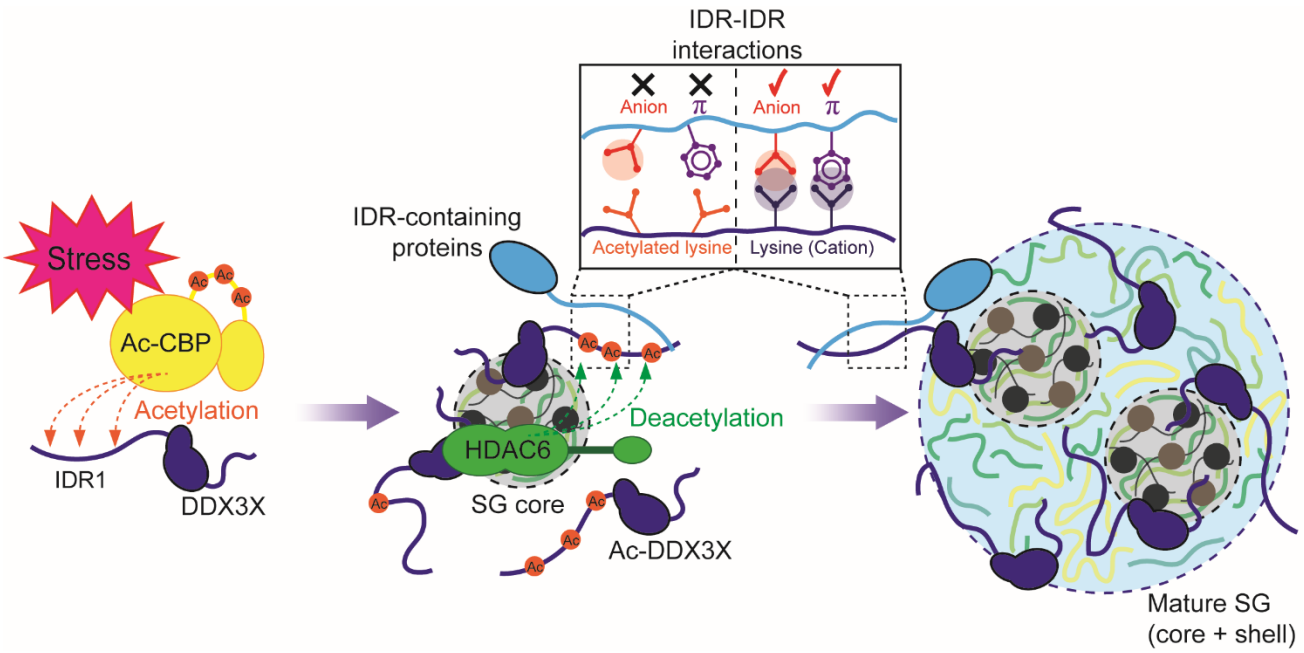


Figure 19. HDAC6 regulates SG maturation through DDX3X deacetylation. Upon stress, CBP becomes activated and elicits acetylation of DDX3X-IDR1 among other proteins, thus impairing LLPS and leading to formation of small SGs. While CBP does not localize to SGs, HDAC6 is recruited to SGs where it deacetylates the IDR1. Positively charged lysine residues in the IDR1 contribute to cation-anion or cation-π interactions. This promotes LLPS of DDX3X and allows other IDR-containing proteins as interaction partners of DDX3X-IDR1 to engage in the formation of large mature SGs.

3.3 Future of acetylation-modulated LLPS

3.3.1 HDAC6 at the crossroad between liquid-liquid and liquid-solid phase separation

Liquid-solid phase transition in biology

Membrane-less organelles generally possess liquid-like properties, but some have more solid-like properties. Importantly, physical properties of phase-separated droplets can change over time *in vitro* and *in vivo*. Some IDR-containing proteins initially form fluid droplets, but become more viscoelastic over time, and eventually stop to exchange molecules with the surroundings and behave as solid (also known as "maturation" or "aging") (Lin et al., 2015; Patel et al., 2015). The material properties of the mature solid state observed *in vitro* are currently unclear, but one possible type is gel. The ALS-related proteins such as FUS and hnRNPA1 can form hydrogels under high-concentration and low-temperature conditions (Kato et al., 2012; Molliex et al., 2015; Murakami et al., 2015) (Fig. 20).

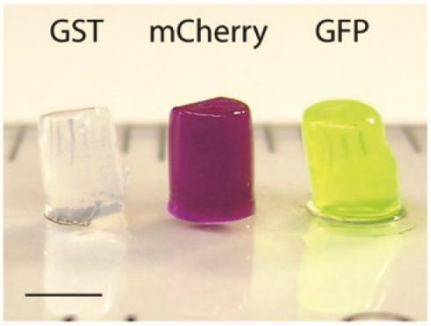


Figure 20. Hydrogels formed by the LCS of FUS. Purified N-terminal LCS of FUS linked to three different protein tags (GST, mCherry, and GFP) adopted a gel-like state when stored at low temperature. The hydrogels retained the cylindrical shape of the silicon tube and exhibit the respective colors of the tags (GST, clear; mCherry, red; GFP, green). Scale bar, 3 mm. Adapted from (Kato et al., 2012).

There were several attempts to reveal the unique properties of such hydrogels; electron microscopy revealed that the FUS-LCS hydrogel contains fibrils, X-ray diffraction uncovered a cross-β pattern reminiscent of amyloids (Kato et al., 2012), and solid state NMR elucidated that 57 of the 214 residues

of FUS LCS form an ordered core, with the remaining residues dynamically disordered (Murray et al., 2017). However, such amyloid-like fibrils are sensitive to heat and SDS, while classical amyloid fibrils resist to boiling and SDS-denaturation. This relative labile character of the hydrogel had suggested that the structure inside is different from amyloids. The atomic structure of segments from several LCSs in hydrogel demonstrated that the short peptides consistently form kinked β-sheet structures. Although cross-β strand formations among side chains give amyloid fibrils their characteristic strength and irreversible nature, this kinked nature of the structure called LARKS (Low-complexity Aromatic-Rich Kinked Segments) prevents strong cross-β strand formation, and contributes to its reversible meshworks (Hughes et al., 2018). Moreover, computational 3D profiling to identify potential LARKS in the human proteome revealed hundreds of human proteins with the sequence propensity to form LARKS. They comprise prominent components of each membrane-less organelle (e.g. DDX3X in SGs, LSM14 in P-bodies and Fibrillarin in nucleolus) and nucleoporins such as Nup54 and Nup98. Purified FG repeats originated from nucleoporins were first reported to form hydrogels *in vitro*, which evoke sieve-like meshwork structure of the nuclear pore complex to achieve selective shuttling of nuclear transport cargos (Frey et al., 2006; Schmidt and Gorlich, 2016). Thus, the affinity and kinetics defined by LARKS may impact on the process where liquid-droplets age and evolve into a solid-like state. It is important to note here that maturation is not observed in droplets formed by interactions among modular domains and their ligands such as SH2+phosphotyrosine and SIM+SUMO; dynamic behaviors remain constant and are determined by the affinity and kinetics of the domain-ligand interactions.

Given that some membrane-less organelles exhibit solid-like properties, and that pathological mutations affect *in vitro* maturation process (Patel et al., 2015) to cause aggregates *in vivo* (Kwiatkowski et al., 2009; Vance et al., 2009), maturation is also likely to occur *in vivo* and plays a key role for physiology and pathology. Balbiani bodies are solid-like membrane-less organelles in immature oocytes of all animals, and are thought to protect the quality of mitochondria and other organelles during long periods of oocyte dormancy (Boke et al., 2016). Purified Xvelo, a prominent

component of Balbiani bodies, forms solid-like structures *in vitro*, which are distinct from liquid-like structures formed by FUS or hnRNPA1. However, whether Xvelo forms solid-like Balbiani bodies *de novo* or they mature from a more liquid-like state is not known, and there are still technical challenging points to monitor the material properties over time *in vivo*. On the other hand, it has become apparent that disease-related mutations can change the threshold for the proteins to form pathogenic solid-like aggregations. Clues regarding the transition from physiological liquids to pathological solids come from recent study on expanded polyQ tracts of huntingtin (HTT) relevant to pathogenic aggregations observed in Huntington's disease (Peskett et al., 2018). HTT exon1 with sub-toxic polyQ lengths (n = 25) can form liquid-like droplets in cells, but does not form irreversible solid-like aggregates. The conversion occurs only when polyQ length extends beyond the threshold for HD ($n \geq 42$ in human). However, sub-toxic polyQ droplets convert to irreversible solid-like structures *in vitro*, suggesting that other cellular components such as chaperones may prevent the liquid-solid phase transition *in vivo*. Nevertheless, this correlation between mutation degree and material property of polyQ indicates that their liquid state is a preliminary step before their "solid" state. Complicatedly, however, liquid state can also be regarded as a safe reservoir to prevent proteins from being aggregated. Phase separation of TDP-43 is triggered by its α-helix, and ALS-associated mutations in that region significantly disrupt LLPS, favoring the conversion to aggregates (Conicella et al., 2016). In sum, it still remains unknown which parameters govern the phase transition kinetics *in vivo*, and several experimental observations, including LLPS, hydrogels *in vitro*, amyloid-like fibers *in vivo*, and more general "aggregate" composed of misfolded proteins all may contribute partly.

There are a few reports that delineated how the material properties of cellular granules are defined and how they are changed over time. In order to identify a sequence-encoded molecular "grammar" underlying the driving forces of phase separation, extensive mutagenesis was carried out for FUS and related proteins (Wang et al., 2018b). Multivalent interactions among tyrosine residues in prion-like domains and arginine residues in RNA-binding domains provide the primary driving force for LLPS. Material properties are modified by other characteristic amino acids; glycine residues

enhance the fluidity, whereas glutamine and serine residues promote hardening. One recent study enabled to control phase transitions spatiotemporally in living cells by using optogenetic protein constructs fused to various IDRs, including FUS, DDX4, and hnRNPA1 (Shin et al., 2017). The constructs undergo light-activated phase separation and form cytoplasmic and nuclear "optoDroplets". When the system is only moderately supersaturated above the threshold by light with low intensity, the optoDroplets are spherical and exhibit liquid-like behaviors, including fusion and efficient FRAP recovery. However, when sufficient high intensity light is used to drive the system, the resulting condensates become more solid-like, with irregular morphology and incomplete FRAP recovery. This platform could be utilized to understand maturation process of liquid droplets within living cells and describe the biological consequence when a cell harbors condensates with a given material property. The majority of these studies focused on kinetics of a single protein species in the experimental system, and no studies mimicked the real situation where membrane-less organelles are composed of many hundreds of proteins. Detailed experimental studies with "artificial organelles" composed of a set of purified proteins including IDR-containing proteins, RNA-binding proteins and chaperons *in vitro* will be required to understand material properties of cellular condensates in future.

HDAC6 as a "double-barreled" protein quality regulator

The most cryptic thing about HDAC6 biology could be the co-existence of two functionally unrelated domains, that is, deacetylase domains and ZnF-UBP in a single molecule. There are a few reports describing the interplay between these domains; for example, K63-linked C-terminal free unanchored ubiquitin chains bind to ZnF-UBP of HDAC6, resulting in increased HDAC6 deacetylase activity (Hao et al., 2013). Yet, we still do not have any rational explanation to connect this interesting domain organization of HDAC6 to its biological function. Considering our novel findings and the discussion so far, we would like to propose here a new concept that HDAC6 is a "double-barreled" protein quality regulator (Fig. 21).

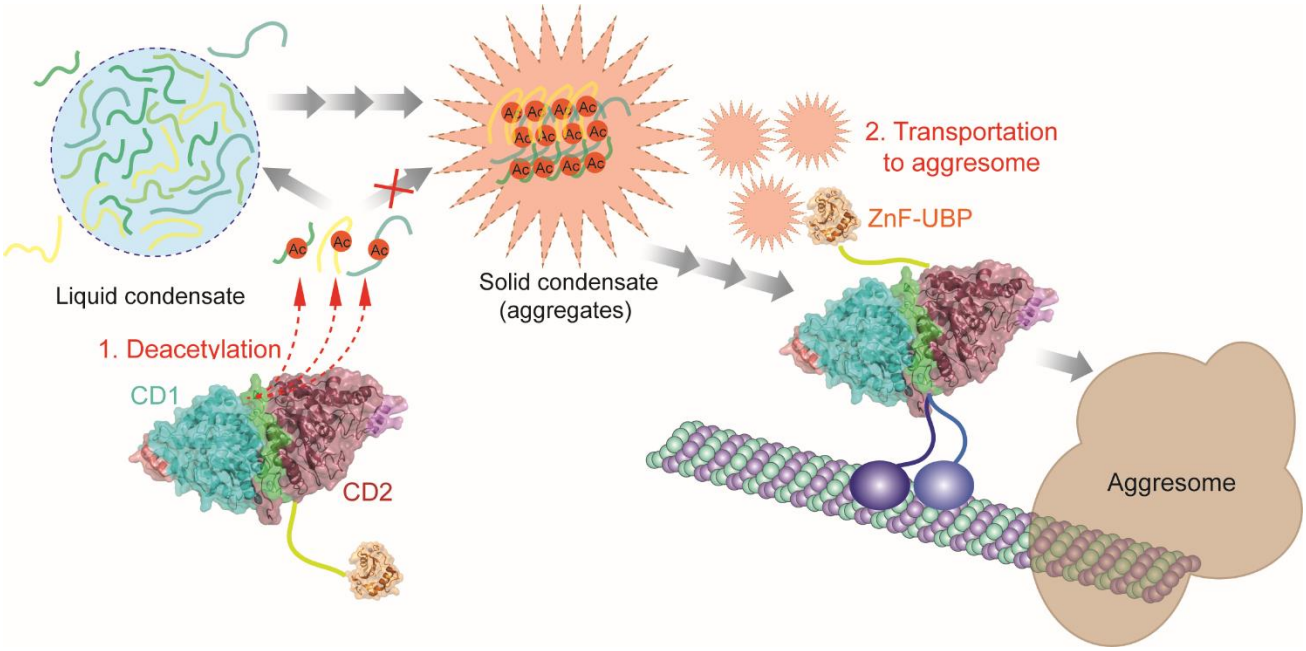


Figure 21. HDAC6 is a "double-barreled" protein quality regulator. HDAC6 removes acetylation so that IDR-containing proteins can appropriately separate into "liquid" phase. In parallel, HDAC6 brings unpreventable aggregates to aggresome for autophagy-mediated clearance.

Given that glutamines in IDRs promote hardening, acetylated lysines may also exhibit similar properties due to their similarity of charges and chemical structures. Thus, the deacetylase domains – the first barrel – support appropriate "liquid-liquid" phase separation by converting acetylated lysine to lysine in IDRs. Enhanced LLPS by HDAC6 may be a way to keep proteins in a healthy "liquid" state after their phase separation, and prevent further droplet ageing and protein aggregation. Moreover, HDAC6 has a "contingency plan" to cope with protein aggregates; ZnF-UBP – the second barrel – binds to misfolded ubiquitinated proteins, and brings them to aggresome for autophagy-mediated clearance. Both functional domains of HDAC6 contribute to prevent aberrant protein aggregation at different scales in our novel concept. Several points should be experimentally tested to confirm this concept, for example:

(1) Clarify the effect of acetylated lysines in liquid droplets

Acetylation inhibits DDX3X LLPS in our study, suggesting that positively charged lysines in DDX3X-IDR1 provide the driving force for its LLPS. Is this effect generalized for all IDRs? Or are there any examples that acetylated lysines contribute to LLPS as glutamine do?

(2) Monitor the metastability from liquid-like SGs to solid-like aggresome over time

Proteasome inhibitors including MG132 are often utilized to induce aggresome in cells (Kawaguchi et al., 2003), but SGs are also observed before aggresome form (Mazroui et al., 2007). How can we integrate mechanisms for these two types of membrane-less organelle formation?

(3) Connect the acetylation status of the protein to its localization

Some proteins such as TDP-43 (Arai et al., 2006; Dewey et al., 2011) and C9orf72 (Boeynaems et al., 2017; Guo et al., 2018b) are components of both SGs and aggresome. With such proteins, can we link the acetylation status of their IDR to their localization? TDP-43 has been reported as an HDAC6 substrate, and indeed deacetylation is required to prevent its aggregation (Cohen et al., 2015), although the acetylation site is not in its IDR.

We hope that our daring concept could be confirmed and/or refined through such analyses, to better explain why HDAC6 has this unique domain organization.

3.3.2 Phase separation in transcription control: beyond granule formation

Our finding about membrane-less organelle formation regulated by (de)acetylation-modulated LLPS is a fundamental mechanism that would explain various cellular processes. This would be a good match especially with emerging concept that transcription is regulated by LLPS (Fig. 22).

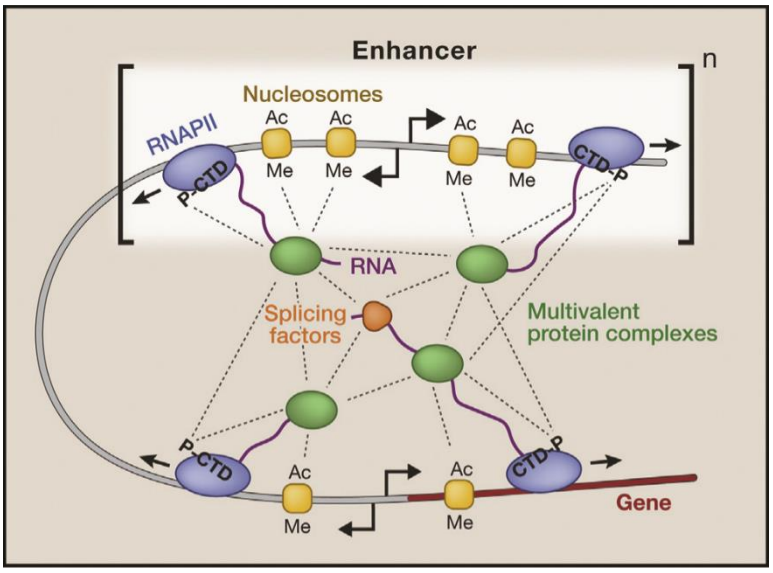


Figure 22. Phase Separation model for transcription. Some of the candidate transcriptional regulators forming the complex are highlighted. P-CTD indicates the phosphorylated CTD of RNA Pol II. PTMs of nucleosomes (acetylation, Ac; methylation, Me) are also highlighted. Divergent transcription at enhancers and promoters produces nascent RNAs. These components associate each other with weak interaction (dashed lines), contributing LLPS. Adapted from (Hnisz et al., 2017).

Three longstanding observations propelled to describe transcription in terms of LLPS. First clue is the observation that labelled nascent transcripts are clustered in the nucleus, which indicates the existence of nuclear site with highly concentrated RNA polymerase II (Pol II) and related transcription machinery. Such a site is now well-known as a “transcription factory” (Iborra et al., 1996; Papantonis and Cook, 2013), which is compatible with other membrane-less organelles observed by microscopy. The structure of transcription factors (TFs) provides us with a second clue. As the history of transcription study narrates it, TFs have been cloned usually based on their structured DNA-binding domains, and thus these domains have been at the center of transcription research. At the same time, however, TFs have a general feature in that they usually harbor IDRs in addition to their DNA-binding domains, and these often function as transcription activation domains in functional assays (Staby et al., 2017). Furthermore, the last clue to engage transcription in the LLPS concept is associated with the unique structure of Pol II.

Transactivation domains (TADs) are functionally defined based on their capacity to activate transcription when tethered to the DNA, as monitored in a suitable assay. They are usually unstructured and form a scaffold harboring binding sites for other transcription coregulators. Because TADs lack sequence similarity and exhibit high degree of LCS, they have been classified according to their amino acid profile: for example, acidic, glutamine-rich, or proline-rich domains have been known from the early days of transcription study (Mitchell and Tjian, 1989) (Fig. 23).

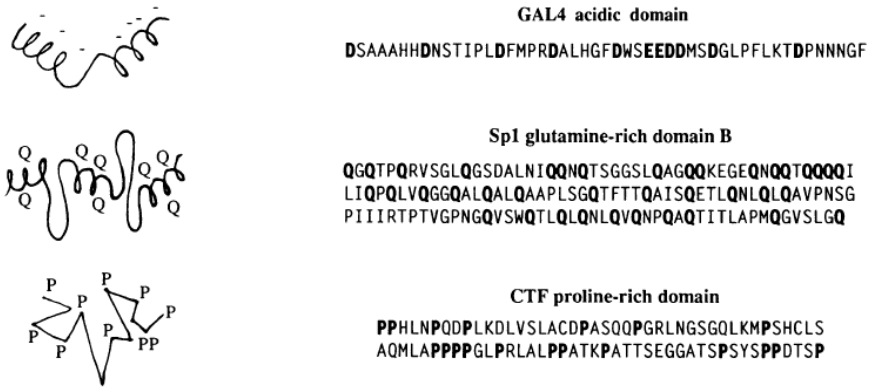


Figure 23. Three major types of transactivation domains (TADs) found in transcription factors. This figure is adapted from the review article published ca. 30 years ago (Mitchell and Tjian, 1989), and this drawing style clearly shows that TADs were already regarded as IDR at that time, but this was not recognized as important then. Amino acid components that characterize the three domain types: acidic, glutamine-rich and proline-rich are written in bold.

As discussed below, some TFs with LCS functioning as TAD have recently been shown to form biomolecular condensates at nuclear loci where target DNA array localizes, suggesting that TADs undergo LLPS to form transcriptional hubs. Thus, TAD-IDR can be regarded as LCS harboring LLPS propensity, but at the same time, it often contains the region functioning as short linear motif (SLiM) and molecular recognition feature (MoRF) (discussed in **1.3.2**) to recruit transcriptional coregulators. This two-faced function is also observed with other IDRs of non-TF proteins such as synapsin1, which can undergo LLPS by itself and also binds to SH3 domain through its PRM domain for efficient LLPS

(discussed in **1.3.2**). Taken together, transcriptional hub formation may be mediated by LLPS in the following way: TFs bind to their target sequence with strong and specific interactions through their DNA binding domain (usually multiple TFs binding in close vicinity), resulting high local concentration of their TADs which can recruit other coregulators through SLiM and/or MoRF and trigger LLPS. It is worth mentioning that this working model is highly similar to the prevalent SG formation "Core first" model (Fig. 24), and that the sole unique point about transcription is how to recruit Pol II; this process is still wrapped in mystery and being extensively studied.

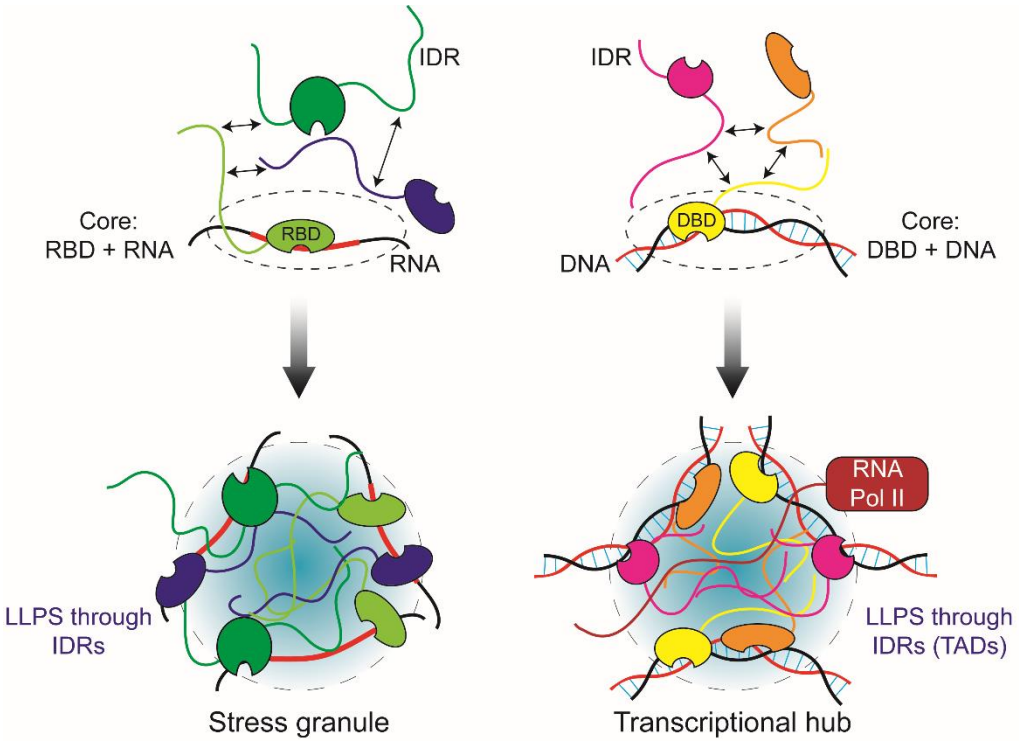


Figure 24. Analogy between formation processes of stress granules (SGs) and transcriptional hubs. RNP components often have both RNA-binding domain (RBD) and IDR. According to a prevalent "Core First" SG formation model, RNP components first form stable core structures through specific, strong interactions between RBD and RNA. Then, the high local concentration of IDRs on SG components would trigger a LLPS to mature SGs. Similarly, TFs are usually equipped with both DNA-binding domain (DBD) and intrinsically disordered TAD. Specific binding between DBD and target DNA sequence cause high local concentration of TADs, resulting in LLPS to establish transcriptional hub. It is still unknown when and how RNA pol II is

recruited to this hub to initiate transcription. Double-headed black arrows indicate the dynamic interaction among IDRs.

Extending from the Pol II catalytic core is a long IDR composed of tandem heptapeptide repeats (52 repeats in human) with consensus "YSPTSPS" sequence which forms the carboxy-terminal domain (CTD) of the largest subunit of RNA polymerase II (RPB1). Although the CTD is not required for the catalytic functions of Pol II, it is essential for several steps in the transcription cycle. The CTD is subjected to multiple PTMs, including phosphorylation, methylation and acetylation (Schroder et al., 2013; Serizawa et al., 1993; Zhao et al., 2016). Very recent studies have established that the CTD can be incorporated into phase-separated liquid droplets and hydrogels *in vitro* (Burke et al., 2015; Kwon et al., 2013; Lu et al., 2018). The LCSs of FUS and its related genes EWS and TAF15 (FET proteins) are translocated onto diverse different DNA-binding domains, which is observed in some cancers including Ewing's sarcoma. These translocated LCSs function as transcriptional activation domains in the context of fusion proteins to regulate transcription for cancer cell survival (Boulay et al., 2017). The purified Pol II CTD directly bound to hydrogel formed by FET proteins, and the degree of binding correlated with the strength of transcriptional activation (Kwon et al., 2013). This binding depends on the length and phosphorylation state of the CTD: a shorter CTD (five heptapeptide repeats) is not binding to the hydrogel, whereas longer ones (ten or more repeats) are efficiently incorporated. In parallel, phosphorylation attenuates the CTD incorporation into the hydrogels, indicating that phosphorylation may inhibit phase separation of the CTD. At the same time, given that FUS has been shown to regulate CTD phosphorylation *in vivo* (Schwartz et al., 2012), it is possible that the CTD and other factors may undergo phase separation in a way that requires CTD phosphorylation. Actually, the CTD phosphorylation seems to either promote or prevent phase separation, depending on its interaction partners. Recent study proposed that CTD itself doen't phase separate and phosphorylation is required for its phase separation. IDR of cyclin T1, a component of positive transcription elongation factor b (P-TEFb) complex known to stimulate Pol II through phosphorylation, undergoes LLPS driven

by its histidine-rich sequence, and recruitment of Pol II into the cyclinT1-IDR droplets is enhanced by its CTD phosphorylation (Lu et al., 2018). DYRK1A, another kinase for CTD (Di Vona et al., 2015), also has histidine-rich sequence, indicating that these key initiation and elongation factors can phase separate to actively recruit Pol II through multivalent interactions to their droplets. Clearly, more studies are required to delineate whole molecular processes from transcription initiation to termination; nevertheless, these observations certainly demonstrate that the phase separation ability of CTD is important for transcription.

The endeavor to clarify the mechanism of transcription with LLPS has just started with detailed analysis of its prominent component Pol II and its regulatory factors mentioned above. Various high-end imaging methods (e.g. lattice light-sheet microscope) in live cells have recently shed new insights into the nature of LCS interactions and their role in transcription (Chong et al., 2018). Transactivation domains form local high-concentration hubs of TFs through their dynamic LCS-LCS interactions. The hubs stabilize DNA binding, recruit Pol II, and activate transcription. LCS-LCS interactions are sequence specific; heterotypic-interactions between LCSs derived from same TF family proteins (e.g. FUS-EWS, both are FET proteins) are stronger than the interactions between different TF family proteins (e.g. FUS-Sp1). Both the composition of LCSs in hubs and their interaction specificity could influence their LLPS propensity and subsequent transcription. It was also shown that LCS-LCS interactions play a role in dysregulated gene expression by the EWS fusion protein associated with Ewing's sarcoma, suggesting that LCS-LCS interactions may represent a new class of therapeutic targets. In addition to TFs, transcription is skillfully regulated by other components such as enhancers and mediators in case of eukaryotic cells (Levine et al., 2014). Enhancers are gene regulatory elements bound by the transcription apparatus and transcription factors, they control cell type-specific gene expression. Super-enhancer (SE) is defined as a cluster of enhancers, which is occupied by high densities of transcriptional machinery. This high density assembly at SEs show to exhibit sharp transitions of formation and dissolution, implying that the high concentrated biomolecules at SEs may be assembled through phase separation (Hnisz et al., 2017). Transcriptional

coactivators BRD4 and MED1 are important components of SEs, and recent studies clarified their contribution to LLPS in detail (Cho et al., 2018; Sabari et al., 2018). BRD4 and MED1 form nuclear puncta, which are also positive for specific transcripts from SEs (Sabari et al., 2018). A companion work showed that Pol II clusters colocalize with this mediator puncta (Cho et al., 2018). These nuclear puncta show liquid-like properties, and are indeed dissolved by 1,6-hexanediol, a compound known to disrupt liquid-like condensates (Lin et al., 2016). This disruption is accompanied with the downregulation of the SE-driven gene expression. Both BRD4 and MED1 contain IDRs and they can phase separate *in vitro*. Droplets of MED1-IDR, which is enriched in conserved serine residues important for its LLPS, incorporate both BRD4 and subunit of RNA pol II (RPB1), suggesting that it can sequester key components of the transcription apparatus. These results strongly support a model (Hnisz et al., 2017) in which transcriptional coactivators form phase-separated compartments to concentrate the transcription apparatus at SEs.

Transcriptional hubs are composed of multiple components, many of which carry PTMs, and interactions between them may provide the valency necessary for LLPS. Given that HDAC6 regulates SG formation by promoting LLPS, it is tempting to speculate that other HDACs in the nucleus may also regulate the formation of transcriptional hubs by modulating LLPS. Further functional analysis about TAD-IDR is awaited to delineate the role of IDR-acetylation in the concept of LLPS-mediated transcription. In particular, our understanding about the molecular grammar of IDRs is not yet sufficient to classify them into subgroups based on their traits so that we can explain the formation of a wide variety of transcriptional hubs enriched with a combination of diverse transcriptional regulators in a unified manner. In the case of DDX3X-IDR1, positively charged lysine residues contribute to LLPS by themselves or with other SG components mainly through cation-anion or cation-π interactions, and thus acetylation works as a negative regulator for LLPS. However, it may also be possible that polar acetylated lysines can promote LLPS, as is the case with polar glutamine in some TADs or prion-like domains. Accordingly, the effect of IDR-acetylation on LLPS would depend on the amino acid composition and sequence of IDR itself and its interaction partners. In order to assess such

context-dependent physicochemical properties of acetylated lysines in LLPS, recombinant proteins with site-specific acetylations would be helpful (Neumann et al., 2008). At the same time, functional hierarchy of substructures within IDR (SLiM, MoRF and LCS) should be also defined in more in detail; this is especially important because it is necessary to consider that acetylated lysines in SLiM or MoRF may also provide a binding platform for acetyl-lysine "reader" domains including BRD mentioned above, and that their multivalent interactions may themselves promote LLPS. This could therefore also be another mechanism for acetylation-modulated LLPS, distinct from our findings that acetylation impairs LLPS through neutralization of IDRs. Given that acetyl-lysine reader proteins are more enriched in the nucleus than the cytosol (Filippakopoulos and Knapp, 2012), it may be possible that acetylation inhibits LLPS in cytosol and enhances it in nucleus (Fig. 25).

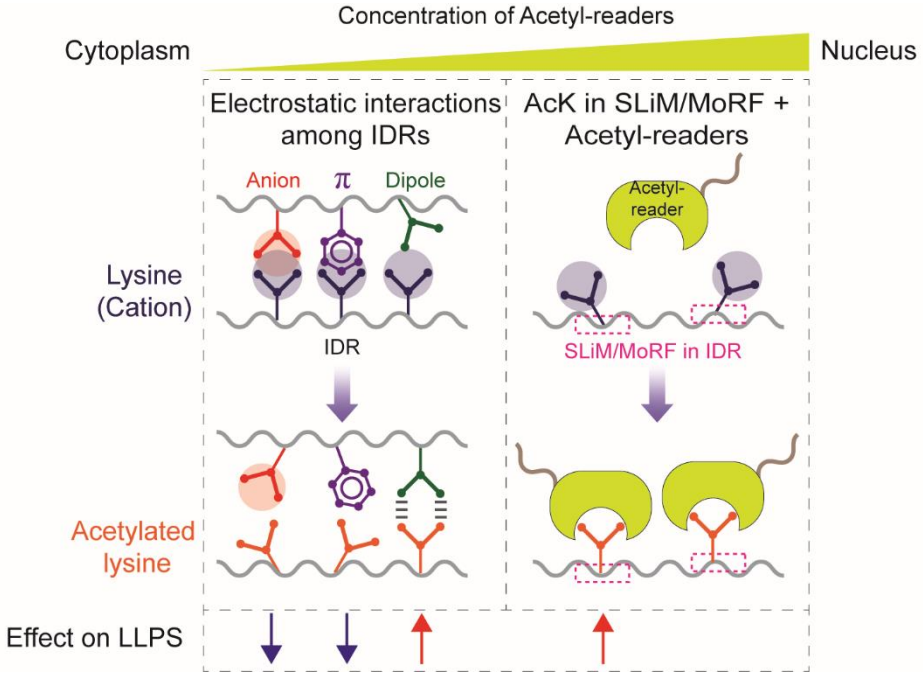


Figure 25. Acetylation-modulated LLPS. Acetylation of IDR impairs the interaction between anionic residues (red) and pi electrons on aromatic residues (purple) in other IDRs, and therefore inhibits LLPS (shown by blue arrows), as observed in DDX3X-IDR1. On the other hand, resulting acetylated lysine (orange) prefers to bind other types of amino acids including polar amino acids (G, Q, N and S; green) and may enhance LLPS. Apart from IDR-IDR interactions, acetylated lysines in SLiM or MoRF of IDRs work as ligands for acetyl-

lysine reader proteins (Acetyl-reader), and the resulting multivalent interactions between acetylated IDR and acetyl-readers would also enhance LLPS (shown by red arrows). This may regulate formation of nuclear biomolecular condensates including transcriptional hubs because acetyl-readers such as BRD family proteins are enriched in the nucleus.

Taken together, functional hierarchy of IDR-substructures for LLPS may also depend on surrounding protein species in addition to their positioning within IDR. Screening with deeply mutated IDRs (Fowler and Fields, 2014) or synthetic IDR libraries of *de novo* amino acid sequences (Simon et al., 2017), which should be supported by machine learning algorithms (Ravarani et al., 2018), would help to systematically delineate the molecular grammar of IDRs including TADs in LLPS. Collectively, we discovered that histone deacetylases regulate LLPS in membrane-less organelle formation, and anticipate that additional biological processes regulated by acetylation-modulated LLPS will be identified in the future.

Materials and Methods

DNA constructs and antibodies

DDX3X, Pex5 and Serbp1 cDNAs were cloned from WT MEF cells into expression vectors (pcDNA3.1 for transient expression or pMSCV for stable expression). Expression vectors for HATs (FLAG-PCAF, HA-CBP, Myc-Tip60 and HA-p300) were kindly supplied from Dr. Renate Voit (German Cancer Research Center, Heidelberg). To construct expression vectors for functional mutants of HDAC6, DDX3X and Serbp1, a plasmid of interest was amplified with appropriate sets of primers, then the PCR product was self-ligated to obtain a mutated plasmid. Primary antibodies were as follows: anti-α-tubulin (DM1A) (Sigma, T9026), anti-Acetylated Tubulin (6-11B-1) (Sigma, T7451), anti-FLAG (M2) (Sigma, F1804), anti-HDAC6 (ab56926), anti-HA (16B12) (ab130275), anti-c-Myc (9E10) (ab32), anti-Phopho-eIF2α (Ser51) (D9G8) (Cell Signaling Technology (CST) #3398), anti-Histone H3 (D1H2) (CST#4499), anti-Acetyl-CBP (Lys1535)/p300 (Lys1499) (CST#4771), anti-eIF2α (D7D3) (CST#5324), anti-CBP (D6C5) (CST#7389), anti-HDAC6 (D21B10) (CST#7612), anti-DDX3 (D19B4) (CST#8192), anti-TIAR (D32D3) (CST#8509), anti-Acetylated-Lysine (CST#9441), anti-Acetylated-Lysine (Ac-K-103) (CST#9681), anti-Acetylated-Lysine (Ac-K2-100) (CST#9814), anti-PABP1 (clone 10E10) (Millipore#04-1467), anti-Acetylated-Lysine (4G12) (Millipore#05-515), anti-acetyl-Histone H3 (Millipore#06-599), anti-DDX3X (Millipore#09-860), anti-Cortactin(p80/85) (4F11) (Millipore#05-180), anti-Acetylated-Lysine (Immunechem, ICP0380), anti-Acetylated-Lysine (1C6) (Thermo Fisher Scientific, MA1-2021), anti-DDX3X (clone 15D1B11) (BioLegend#658602), anti-Ub (P4D1) (Santa Cruz, sc-8017), anti-[K(Ac)40]-α-tubulin (Enzo, BML-SA452-0100) and anti-G3BP (Aviva Systems Biology, ARP37713_T100). Anti-mouse-HDAC6 was developed in Matthias laboratory, FMI (Banerjee et al., 2014). Secondary antibodies were as follows: Amersham ECL Mouse IgG, HRP-linked whole Ab from sheep (GE Healthcare, NA931V), Amersham ECL Rabbit IgG, HRP-linked whole Ab from donkey (GE Healthcare, NA934V), Alexa Fluor 488 Goat anti-Mouse IgG (H+L) Secondary Antibody (Invitrogen, A11001), Alexa Fluor 568 Goat anti-Mouse IgG (H+L) Secondary Antibody (Invitrogen, A11004), Alexa Fluor 488 Goat anti-Rabbit IgG (H+L) Secondary Antibody (Invitrogen, A11034) and Alexa Fluor 568 Goat anti-Rabbit IgG (H+L)

Secondary Antibody (Invitrogen, A11036).

Establishment and characterization of DDX3X-K118Ac-specific antibody

The DDX3X-K118 acetylated peptide, CDRSGFGK(Ac)FERG (PSL Peptide Speciality Laboratories) was conjugated with mcKLH by Imject Maleimide Activated Carrier Protein Spin Kits (Thermo scientific) and the KLH conjugated peptide was used to immunize two rabbits (Pocono Rabbit Farm & Laboratory). Collected serum was passed over K118 acetylated peptide CDRSGFGK(Ac)FERG-conjugated agarose column prepared by SulfoLink Immobilization Kit for Peptides (Thermo scientific) and eluted with 0.2 M glycine-HCl, pH 2.0 and neutralized. Then, the elution was passed over K118 unacetylated peptide CDRSGFGKFERG-conjugated agarose column and the flow through was used as DDX3X-K118Ac-specific antibody. The specificity for K118Ac was assessed by ELISA using K118 acetylated and unacetylated peptide-coated plates (TaKaRa) and 1-Step Slow TMB-ELISA (Thermo scientific).

Cell culture

MEF cells, HEK293T cells, HeLa cells and Plat-E packaging cells were cultured in DMEM supplemented with 10% heat-inactivated FBS, penicillin, and streptomycin at 37 °C and 5% CO_2.

Transfection

HEK293T cells, HeLa cells and Plat-E packaging cells were transfected with Lipofectamine 2000 (Life Technologies) or FuGENE HD (Promega), and MEF cells were transfected with 4D-Nucleofector™ System (Lonza, Kit: P4 Primary Cell 4D-Nucleofector™ X, Pulse program: CZ-167), following manufacturer's protocol.

Immunoblotting and immunoprecipitation

Cells were washed by ice-cold PBS and lysed in Triton lysis buffer (50 mM Tris, pH 8.0, 150 mM

NaCl, 1 mM EDTA, 0.1% TritonX-100 and Complete EDTA-free protease inhibitors (Roche)) for analysis. To detect protein acetylation, 0.2 μM Trichostatin A and 5 mM nicotinamide was added to PBS for wash, and 10 μM Trichostatin A, 10 mM nicotinamide, 50 mM sodium butyrate were added to Triton lysis buffer. To detect protein phosphorylation, Pierce Protease and Phosphatase Inhibitor Mini Tablets (Thermo scientific) were added to Triton lysis buffer instead of Complete EDTA-free protease inhibitors. Five hundred μg protein lysates were incubated with the specific antibody for overnight at 4 °C, and immunoprecipitated with Dynabeads ProteinG (invitrogen) for 1 h at 4 °C. Samples were boiled for 10 min in SDS-PAGE sample buffer, and separated with 4-12% NuPAGE gels (invitrogen). Proteins were transferred onto PVDF membranes (Immobilon-P, Millipore), proved with specific primary antibody for overnight and secondary antibody for 1 h under 5% non-fat dry milk in TBS or 5% BSA, 0.1% Tween20 in TBS blocking conditions and detected with Amersham ECL Western blotting reagent (GE Healthcare). For inhibition of HDACs in cells, concentration of each HDAC inhibitor was following: 20 mM Nicotinamide (Sigma), 5 μM MS275 (Sigma), 2 μM SAHA (Sigma), 2 μM TSA (MBL), 50 μM Bufexamac (Sigma, European Pharmacopoeia (EP) Reference Standard), 10 μM Tubacin (Sigma), 10 μM Tubastatin A (Sigma), and 5 mM Sodium butyrate (Sigma), respectively. Cells were harvested for lysis after 16 h treatment with these inhibitors.

Identification of the DDX3X acetylation sites

The protein pellets were reduced with 10 mM TCEP, alkylated with 20 mM iodoacetamide and cleaved first with 0.1 μg Lys-C (Wako) for 6 h and then with 0.1 μg porcine sequencing grade trypsin (Promega) for overnight. The extracted peptides were analyzed by capillary liquid chromatography tandem mass spectrometry with an EASY-nLC 1000 using the two-column set up (Thermo Scientific). The peptides were loaded in 0.1% formic acid, 2% acetonitrile in H_2O onto a peptide trap (Acclaim PepMap 100, 75 μm x 2 cm, C18, 3 μm, 100 Å) at a constant pressure of 600 bar. Then they were separated, at a flow rate of 150 nl/min with a linear gradient of 2-6% buffer B in buffer A in 3 min followed by a linear increase from 6 to 22% in 40 min, 22-28% in 9 min, 28-36% in 8 min, 36-80% in

1 min and 12 min at 80% buffer B in buffer A (buffer A: 0.1% formic acid, buffer B: 0.1% formic acid in acetonitrile) on a 50 μm x 15 cm ES801 C18, 2 μm, 100Å column mounted on a DPV ion source (New Objective) connected to a Orbitrap Fusion (Thermo Scientific). The data were acquired using 120000 resolution for the peptide measurements in the Orbitrap and a top T (3 s) method with HCD fragmentation for each precursor and fragment measurement in the linear ion trap. MASCOT 2.5 was used in the Decoy mode to search the Swiss-Prot mouse version 2015_01 including common contaminants. The enzyme specificity was set to trypsin allowing for up to three incomplete cleavage sites. Carbamidomethylation of cysteine (+57.0245 Da) was set as a fixed modification, oxidation of methionine (+15.9949 Da) and acetylation of lysines (+42.0106 Da) were set as variable modifications. Parent ion mass tolerance was set to 10 ppm and fragment ion mass tolerance to 0.6 Da. The results were validated with the program Scaffold (version: 4.4) and Scaffold PTM (version: 2.2) (Proteome Software, Portland, USA). Peptide identifications were accepted if they could be established at greater than 0.1% probability as calculated in Scaffold and acetylation sites were accepted if they had a greater than 80% site probability as calculated with Scaffold PTM.

Interaction proteomics

FLAG-DDX3X was immunoprecipitated with Anti-FLAG M2 Magnetic Beads (Sigma) from cellular lysate of each 293T sample expressing either DDX3X WT or mutants. Cellular lysate of parental 293T cells was used as a negative control. Precipitants were digested on beads by Lys-C (Wako) in digestion buffer (3 M GuHCl, 20 mM EPPS, pH 8.5, 10 mM CAA and 5 mM TCEP), further digested by trypsin (Promega), and analyzed by LC-MS/MS, essentially as described (Ostapcuk et al., 2018). In short, the peptides were separated with an EASY-nLC 1000 on a 50 μm × 15 cm ES801 C18, 2 μm, 100 Å column (Thermo Scientific) mounted on a DPV ion source (New Objective). They were measured with a Orbitrap Fusion (Thermo Scientific) using a top T (3 s) method as recommended by the manufacturer (Thermo Scientific). Andromeda implemented in MaxQuant (version: 1.5.3.8) (Cox et al., 2011) was used to search the mouse subset of the UniProt (version: 2017_04) combined with the contaminant

database from MaxQuant and label-free quantification (LFQ (Cox et al., 2014)) was used with a protein and peptide FDR of 0.01. Statistical analysis was done in Perseus (version: 1.5.2.6) (Hubner et al., 2010; Tyanova et al., 2016). Results were filtered to remove reverse hits, contaminants and peptides found in only one sample. Missing values were imputed and potential interactors were determined using Student's two-tailed t-test and visualized by a volcano plot. Significance lines corresponding to a FDR of 0.03 and a S0 of 1.5 are shown in the corresponding figures. Results were exported from Perseus and visualized using statistical computing language R.

Biotinylated isoxazole (b-isox) -mediated precipitation

Cells were washed by ice-cold PBS and lysed in EE buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM EDTA, 2.5 mM EGTA, 0.1% NP-40, 10% Glycerol, 1 μM DTT and Complete EDTA-free protease inhibitors (Roche)) To detect DDX3X acetylation, 10 μM Tubacin was added to EE buffer. Biotinylated isoxazole (Sigma) in DMSO was added to cell lysates at 100 μM final concentration. The reaction solutions were incubated for 4 h at 4 °C and then centrifuged at 13200 rpm for 10 min. Precipitates were washed five times in EE buffer before SDS solubilization, following analyzed by immunoblotting as mentioned above.

Immunofluorescence microscopy and live-cell imaging

Stress granules were induced by the following conditions: 1 mM sodium arsenite (Millipore) for 15 min-1 h, 2 mM H_2O_2 (Sigma) for 1 h, 3 mM Diethyl maleate (Sigma) for 1 h, 20 μM CCCP (Sigma) in glucose-free DMEM (Gibco) with 10% FBS for 1 h, 40 μM clotrimazole (Sigma) in glucose-free DMEM with 10% FBS for 3 h, 20 μg/ml puromycin (Sigma) for 3 h, 10 μM thapsigargin (Invitrogen) for 1 h, 10 μM MG132 (Sigma) for 3 h, heat shock at 43 °C for 1 h and 0.5 M sorbitol (Sigma) for 1 h. Catalytic activities of HDAC6 or CBP were inhibited by treatment of 10 μM Tubacin (Sigma) or 10 μM A-485 (Lucerna-Chem) for 3 h, respectively. For knockdown of CBP, ON-TARGETplus SMARTpool siRNA mouse CBP (Dharmacon) were delivered via nucleofection (2 μM). Cells on a

coverglass were washed with ice-cold PBS, then fixed with ice-cold methanol or 4% paraformaldehyde. After permeabilization with 0.5% TritonX-100 or 0.005% digitonin in PBS for 10 min, the cells were incubated with specific primary and secondary antibodies in 10% goat serum (Sigma) blocking buffer, then mounted with ProLong Gold Antifade Reagent with DAPI (Cell Signaling Technology #8961). Images were captured by Axioimager Z1 microscope (Zeiss). SGs in images were defined by a plugin of Fiji (SG counter). Volume of SG $V = 4\pi R^3/3$, where R is SG radius, was calculated for each SG, and total volume of SGs in a cell was determined. In each violin plot, the thin line at the center represents the 95% confidence interval, and the bottom and top of the box represent the 25th and 75th percentile of the data, with the median value represented by a horizontal bold line. On each side of the thin line is a kernel density estimation showing the distribution shape of the data points from minima to maxima. For live-cell imaging, WT or DDX3X KO MEF cells were transfected with expression vectors encoding mCherry-DDX3X WT or mutants, cultured for a day on a glass bottom dish 35 mm (ibidi), and images were captured with a spinning-disk confocal microscope (Olympus IX81 + Yokogawa CSU-X1 scan head) with a 561 nm laser for mCherry signals.

Mathematical modelling of SG growth

Live-cell image stacks, captured as mentioned above, were subjected to maximum intensity projection, and SGs were defined by SG counter. The time point when cells are treated with arsenite was set as t = 0, and the initiation time of SGs (T) was defined as the time point when SG is observed for the first time. The obtained total SG area is regressed on Avrami equation:

$$A(t)=A_F(1-e^{-k(t-T)})$$

and two parameters A_F and k were estimated from this fit. To simulate kinetics of each DDX3X construct, the effect of size scaling was adjusted: the initial area A_0 was measured at t = 0 for each cell, and the scaling ratio r (= A_F / A_0) was obtained. The best estimate and associated error (SD) for k and r were calculated from the fit (n = 5 for each construct). To visualize the data, the mean value of r for WT is arbitrarily set as 100, and the curves obtained with the best estimates of T, k and r were shown.

The areas where each curve exists with three parameters: T, k and r in the range of mean ± SD were also shown.

Establishment of KO cell lines using CRISPR/Cas9 genome editing

HDAC6 KO HEK293T cells were established by CRISPR/Cas9 genome editing. The guide sequence targeting human HDAC6 was designed by the CRISPR design tool at http://www.genome-engineering.org/crispr/ and cloned into pX330-Cas9-T2A-mCherry vector (Addgene). The inserted 20 nucleotides is following:

5'-GGTGGAATCCTGGCCGGTTG-3'

The pX330 vector was transfected into HEK293T cells as mentioned above. Two days after transfection, the mCherry positive cells were collected by FACS. Then single cell clones were sorted again week after and expanded. Screening for HDAC6 knockout was done by immunoblotting. Genomic DNA was purified from the HDAC6 KO clone and the region surrounding PAM of the sgRNA was cloned into pGEM-T Easy vector (Promega) for sequencing after amplification with a pair of primers:

5'-GGCAGAGAGGTGGGGTCCTC-3'

5'-CTCATCAGGGCGGGTCTAGG-3'

To determine the indels of individual alleles, the amplicons from 20 bacterial colonies were sequenced. Establishment of DDX3X KO MEFs followed the same procedure. The inserted nucleotides and a pair of primers for genotyping are followings:

5'-TGGCAGTGGAAAATGCGCT -3' (for guide RNA)

5'-GAGGAGGGCACACGTCTTAC-3' (for genotyping)

5'-AACTTAAAGAGCTGCGCCAC-3' (for genotyping)

Protein expression and purification

The DNA fragments of mouse DDX3X-IDR1 (a.a. 2-167) and DDX3X-full-length (a.a. 2-662) were

inserted into the BamHI sites of the pET28 vector. All proteins were expressed in *E. Coli* BL21 (DE3) cells with 1.0 mM IPTG at 16 °C for overnight. Cells were lysed with a lysis buffer containing 20 mM Tris, pH 7.5, 500 mM NaCl, 2 mM DTT, Complete EDTA-free protease inhibitors (Roche) for 30 min on ice, followed by sonication. The lysate was centrifuged at 15000 rpm for 30 min. The supernatant was incubated with Ni-NTA resin (QIAGEN) for 1 h at 4 °C. The Ni-NTA resin was packed in a column and washed with a buffer containing 20 mM Tris, pH 7.5, 500 mM NaCl, 20 mM imidazole, 2 mM DTT. The bound proteins were eluted with a buffer containing 20 mM Tris, pH 7.5, 500 mM NaCl, 500 mM imidazole, 2 mM DTT. The purified His-SUMO-tagged proteins were concentrated with Amicon Ultra centrifugal filters (Millipore), and digested by SUMO protease (MCLAB) for overnight at 4 °C. Then, the digested samples were passed over Ni-NTA resin and the flow through was collected and concentrated. They were subjected to size exclusion chromatography on a Superdex 200 Increase 10/300 GL column (GE Healthcare) equilibrated in the Gel Filtration buffer, 20 mM Tris, pH 7.5, 250 mM NaCl, and 2 mM DTT. The fractions were concentrated to 20 mg/ml and stored at -80 °C. For full-length protein, 1 M NaCl buffers were used through the whole purification process. The purities of the purified proteins were tested on SDS-PAGE gels, and the concentrations were determined by absorbance at UV at 280 nm.

In vitro droplet formation assay

For the phase separation of purified DDX3X, experiment was performed in 20 mM Tris, pH 7.5, 2 mM DTT. The concentrations of DDX3X and NaCl are 0-200 μM and 0-250 mM, respectively as indicated in the figure legends. As a molecular crowder, PEG4000 (Sigma) was added to samples at 10% (w/v). For acetylation of DDX3X-IDR1 *in vitro*, 40 μg of the IDR1 was incubated with 2.5 μg of human recombinant CBP catalytic domain (Enzo) and 10 mM acetyl-CoA (Sigma) for 3 h at 37 °C. For deacetylation of DDX3X-IDR1 *in vitro*, *D. rerio* HDAC6 CD1-CD2 (a.a.40-831) was expressed in Sf9 insect cells, purified based on the method previously described (buffer and salt were adjusted to the droplet formation assay), and added into DDX3X-IDR1 acetylation reaction mixture. Alexa Fluor 488

Protein Labeling Kit (Molecular Probes) was used for fluorescent labeling of the IDRs, followed by dialysis against Gel Filtration buffer using Slide-A-Lyzer™ MINI Dialysis Devices (Thermo scientific). The droplets were observed using an Axioimager Z1 microscope (Zeiss) and Spinning-Disk Confocal microscope, and turbidity of the samples (OD600) is quantified with NanoDrop (Thermo scientific). For temperature-dependent droplet dissolution and condensation experiments, samples were mounted inside small compartments sealed with a pegylated coverslip on a custom-built temperature-controlled stage (Mittasch et al., 2018). For imaging, the heating/cooling system was mounted on a Nikon TiE stand equipped with a Perfect Focus System. Image stacks were captured every 10 s using a 40x air objective (0.95NA) combined with a spinning-disk confocal unit CSU-X1 (Yokogawa) and an EMCCD camera IXON+ (Andor). Image data was analyzed with MATLAB. In brief, image stacks were subjected to maximum intensity projection, the background was subtracted, and they were binarized after further filtering. In the binarized image-series, particle number was analyzed.

Computational prediction of disordered regions

Computational prediction of disordered regions was done with the PONDR VSL2 and VL3-BA programs. As queries for the prediction, protein sequences from the UniProt database released in March 2017 were used. Missing protein entries in this version were cited from the one released in February 2012.

References

Aguzzi, A., and Altmeyer, M. (2016). Phase Separation: Linking Cellular Compartmentalization to Disease. Trends Cell Biol *26*, 547-558.

Allfrey, V.G., Faulkner, R., and Mirsky, A.E. (1964). Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proc Natl Acad Sci U S A *51*, 786-794.

Arai, T., Hasegawa, M., Akiyama, H., Ikeda, K., Nonaka, T., Mori, H., Mann, D., Tsuchiya, K., Yoshida, M., Hashizume, Y.*, et al.* (2006). TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem Biophys Res Commun *351*, 602-611.

Ariumi, Y., Kuroki, M., Abe, K., Dansako, H., Ikeda, M., Wakita, T., and Kato, N. (2007). DDX3 DEAD-box RNA helicase is required for hepatitis C virus RNA replication. J Virol *81*, 13922-13926.

Banani, S.F., Lee, H.O., Hyman, A.A., and Rosen, M.K. (2017). Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol *18*, 285-298.

Banani, S.F., Rice, A.M., Peeples, W.B., Lin, Y., Jain, S., Parker, R., and Rosen, M.K. (2016). Compositional Control of Phase-Separated Cellular Bodies. Cell *166*, 651-663.

Banerjee, I., Miyake, Y., Nobs, S.P., Schneider, C., Horvath, P., Kopf, M., Matthias, P., Helenius, A., and Yamauchi, Y. (2014). Influenza A virus uses the aggresome processing machinery for host cell entry. Science *346*, 473-477.

Bannister, A.J., Miska, E.A., Gorlich, D., and Kouzarides, T. (2000). Acetylation of importin-alpha nuclear import factors by CBP/p300. Curr Biol *10*, 467-470.

Beckham, C., Hilliker, A., Cziko, A.M., Noueiry, A., Ramaswami, M., and Parker, R. (2008). The DEAD-box RNA helicase Ded1p affects and accumulates in Saccharomyces cerevisiae P-bodies. Mol Biol Cell *19*, 984-993.

Bernardi, R., and Pandolfi, P.P. (2007). Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. Nat Rev Mol Cell Biol *8*, 1006-1016.

Bertos, N.R., Gilquin, B., Chan, G.K., Yen, T.J., Khochbin, S., and Yang, X.J. (2004). Role of the tetradecapeptide repeat domain of human histone deacetylase 6 in cytoplasmic retention. J Biol Chem *279*, 48246-48254.

Biggar, K.K., and Li, S.S. (2015). Non-histone protein methylation as a regulator of cellular signalling and function. Nat Rev Mol Cell Biol *16*, 5-17.

Bitler, B.G., Wu, S., Park, P.H., Hai, Y., Aird, K.M., Wang, Y., Zhai, Y., Kossenkov, A.V., Vara-Ailor, A., Rauscher, F.J., III*, et al.* (2017). ARID1A-mutated ovarian cancers depend on HDAC6 activity. Nat Cell Biol *19*, 962-973.

Boeynaems, S., Bogaert, E., Kovacs, D., Konijnenberg, A., Timmerman, E., Volkov, A., Guharoy, M., De Decker, M., Jaspers, T., Ryan, V.H.*, et al.* (2017). Phase Separation of C9orf72 Dipeptide Repeats Perturbs Stress Granule Dynamics. Mol Cell *65*, 1044-1055 e1045.

Boke, E., Ruer, M., Wuhr, M., Coughlin, M., Lemaitre, R., Gygi, S.P., Alberti, S., Drechsel, D., Hyman, A.A., and Mitchison, T.J. (2016). Amyloid-like Self-Assembly of a Cellular Compartment. Cell *166*, 637-650.

Botlagunta, M., Vesuna, F., Mironchik, Y., Raman, A., Lisok, A., Winnard, P., Jr., Mukadam, S., Van Diest, P., Chen, J.H., Farabaugh, P.*, et al.* (2008). Oncogenic role of DDX3 in breast cancer biogenesis. Oncogene *27*, 3912-3922.

Bouchard, J.J., Otero, J.H., Scott, D.C., Szulc, E., Martin, E.W., Sabri, N., Granata, D., Marzahn, M.R., Lindorff-Larsen, K., Salvatella, X.*, et al.* (2018). Cancer Mutations of the Tumor Suppressor SPOP Disrupt the Formation of Active, Phase-Separated Compartments. Mol Cell *72*, 19-36 e18.

Boulay, G., Sandoval, G.J., Riggi, N., Iyer, S., Buisson, R., Naigles, B., Awad, M.E., Rengarajan, S., Volorio, A., McBride, M.J.*, et al.* (2017). Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. Cell *171*, 163-178 e119.

Boyault, C., Gilquin, B., Zhang, Y., Rybin, V., Garman, E., Meyer-Klaucke, W., Matthias, P., Muller, C.W., and Khochbin, S. (2006). HDAC6-p97/VCP controlled polyubiquitin chain turnover. EMBO J *25*, 3357-3366.

Boyault, C., Zhang, Y., Fritah, S., Caron, C., Gilquin, B., Kwon, S.H., Garrido, C., Yao, T.P., Vourc'h, C., Matthias, P.*, et al.* (2007). HDAC6 controls major cell response pathways to cytotoxic accumulation of protein aggregates. Genes Dev *21*, 2172-2181.

Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoege, C., Gharakhani, J., Julicher, F., and Hyman, A.A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science *324*, 1729-1732.

Brangwynne, C.P., Tompa, P., and Pappu, R.V. (2015). Polymer physics of intracellular phase transitions. Nature

Physics *11*, 899-904.

Brownell, J.E., and Allis, C.D. (1995). An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in Tetrahymena macronuclei. Proc Natl Acad Sci U S A *92*, 6364-6368.

Brunet, A., Bonni, A., Zigmond, M.J., Lin, M.Z., Juo, P., Hu, L.S., Anderson, M.J., Arden, K.C., Blenis, J., and Greenberg, M.E. (1999). Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. Cell *96*, 857-868.

Bryce, N.S., Clark, E.S., Leysath, J.L., Currie, J.D., Webb, D.J., and Weaver, A.M. (2005). Cortactin promotes cell motility by enhancing lamellipodial persistence. Curr Biol *15*, 1276-1285.

Buchan, J.R., Kolaitis, R.M., Taylor, J.P., and Parker, R. (2013). Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. Cell *153*, 1461-1474.

Burke, K.A., Janke, A.M., Rhine, C.L., and Fawzi, N.L. (2015). Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. Mol Cell *60*, 231-241.

Busch, D.J., Houser, J.R., Hayden, C.C., Sherman, M.B., Lafer, E.M., and Stachowiak, J.C. (2015). Intrinsically disordered proteins drive membrane curvature. Nat Commun *6*, 7875.

Calabretta, S., and Richard, S. (2015). Emerging Roles of Disordered Sequences in RNA-Binding Proteins. Trends Biochem Sci *40*, 662-672.

Chahar, H.S., Chen, S., and Manjunath, N. (2013). P-body components LSM1, GW182, DDX3, DDX6 and XRN1 are recruited to WNV replication sites and positively regulate viral replication. Virology *436*, 1-7.

Chang, P.C., Chi, C.W., Chau, G.Y., Li, F.Y., Tsai, Y.H., Wu, J.C., and Wu Lee, Y.H. (2006a). DDX3, a DEAD box RNA helicase, is deregulated in hepatitis virus-associated hepatocellular carcinoma and is involved in cell growth control. Oncogene *25*, 1991-2003.

Chang, S., McKinsey, T.A., Zhang, C.L., Richardson, J.A., Hill, J.A., and Olson, E.N. (2004). Histone deacetylases 5 and 9 govern responsiveness of the heart to a subset of stress signals and play redundant roles in heart development. Mol Cell Biol *24*, 8467-8476.

Chang, S., Young, B.D., Li, S., Qi, X., Richardson, J.A., and Olson, E.N. (2006b). Histone deacetylase 7 maintains vascular integrity by repressing matrix metalloproteinase 10. Cell *126*, 321-334.

Chao, C.H., Chen, C.M., Cheng, P.L., Shih, J.W., Tsou, A.P., and Lee, Y.H. (2006). DDX3, a DEAD box RNA helicase with tumor growth-suppressive property and transcriptional regulation activity of the p21waf1/cip1 promoter, is a candidate tumor suppressor. Cancer Res *66*, 6579-6588.

Chen, W.C., McBride, W.H., Iwamoto, K.S., Barber, C.L., Wang, C.C., Oh, Y.T., Liao, Y.P., Hong, J.H., de Vellis, J., and Shau, H. (2002). Induction of radioprotective peroxiredoxin-I by ionizing irradiation. J Neurosci Res *70*, 794-798.

Chen, X., Wong, J.Y., Wong, P., and Radany, E.H. (2011). Low-dose valproic acid enhances radiosensitivity of prostate cancer through acetylated p53-dependent modulation of mitochondrial membrane potential and apoptosis. Mol Cancer Res *9*, 448-461.

Cho, W.K., Spille, J.H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, II (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science *361*, 412-415.

Choi, S.J., Lee, H.C., Kim, J.H., Park, S.Y., Kim, T.H., Lee, W.K., Jang, D.J., Yoon, J.E., Choi, Y.I., Kim, S.*, et al.* (2016). HDAC6 regulates cellular viral RNA sensing by deacetylation of RIG-I. EMBO J *35*, 429-442.

Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G.M., Cattoglio, C., Heckert, A., Banala, S., Lavis, L., Darzacq, X.*, et al.* (2018). Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science *361*.

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V., and Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science *325*, 834-840.

Choudhary, C., Weinert, B.T., Nishida, Y., Verdin, E., and Mann, M. (2014). The growing landscape of lysine acetylation links metabolism and cell signalling. Nat Rev Mol Cell Biol *15*, 536-550.

Clevers, H., and Nusse, R. (2012). Wnt/beta-catenin signaling and disease. Cell *149*, 1192-1205.

Cohen, T.J., Hwang, A.W., Restrepo, C.R., Yuan, C.X., Trojanowski, J.Q., and Lee, V.M. (2015). An acetylation switch controls TDP-43 function and aggregation propensity. Nat Commun *6*, 5845.

Conicella, A.E., Zerze, G.H., Mittal, J., and Fawzi, N.L. (2016). ALS Mutations Disrupt Phase Separation Mediated by alpha-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. Structure *24*, 1537-1549.

Cooper, E.M., Cutcliffe, C., Kristiansen, T.Z., Pandey, A., Pickart, C.M., and Cohen, R.E. (2009). K63-specific deubiquitination by two JAMM/MPN+ complexes: BRISC-associated Brcc36 and proteasomal Poh1. EMBO J *28*, 621-631.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics *13*, 2513-2526.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J Proteome Res *10*, 1794-1805.

Cruciat, C.M., Dolde, C., de Groot, R.E., Ohkawara, B., Reinhard, C., Korswagen, H.C., and Niehrs, C. (2013). RNA helicase DDX3 is a regulatory subunit of casein kinase 1 in Wnt-beta-catenin signaling. Science *339*, 1436-1441.

Dao, T.P., Kolaitis, R.M., Kim, H.J., O'Donovan, K., Martyniak, B., Colicino, E., Hehnly, H., Taylor, J.P., and Castaneda, C.A. (2018). Ubiquitin Modulates Liquid-Liquid Phase Separation of UBQLN2 via Disruption of Multivalent Interactions. Mol Cell *69*, 965-978 e966.

Das, R.K., Ruff, K.M., and Pappu, R.V. (2015). Relating sequence encoded information to form and function of intrinsically disordered proteins. Curr Opin Struct Biol *32*, 102-112.

Datta, S.R., Katsov, A., Hu, L., Petros, A., Fesik, S.W., Yaffe, M.B., and Greenberg, M.E. (2000). 14-3-3 proteins and survival kinases cooperate to inactivate BAD by BH3 domain phosphorylation. Mol Cell *6*, 41-51.

de Zoeten, E.F., Wang, L., Butler, K., Beier, U.H., Akimova, T., Sai, H., Bradner, J.E., Mazitschek, R., Kozikowski, A.P., Matthias, P.*, et al.* (2011). Histone deacetylase 6 and heat shock protein 90 control the functions of Foxp3(+) T-regulatory cells. Mol Cell Biol *31*, 2066-2078.

Delarue, M., Brittingham, G.P., Pfeffer, S., Surovtsev, I.V., Pinglay, S., Kennedy, K.J., Schaffer, M., Gutierrez, J.I., Sang, D., Poterewicz, G.*, et al.* (2018). mTORC1 Controls Phase Separation and the Biophysical Properties of the Cytoplasm by Tuning Crowding. Cell *174*, 338-349 e320.

Dewey, C.M., Cenik, B., Sephton, C.F., Dries, D.R., Mayer, P., 3rd, Good, S.K., Johnson, B.A., Herz, J., and Yu, G. (2011). TDP-43 is directed to stress granules by sorbitol, a novel physiological osmotic and oxidative stressor. Mol Cell Biol *31*, 1098-1108.

Di Vona, C., Bezdan, D., Islam, A.B., Salichs, E., Lopez-Bigas, N., Ossowski, S., and de la Luna, S. (2015).

Chromatin-wide profiling of DYRK1A reveals a role as a gene-specific RNA polymerase II CTD kinase. Mol Cell *57*, 506-520.

Dikic, I., Wakatsuki, S., and Walters, K.J. (2009). Ubiquitin-binding domains - from structures to functions. Nat Rev Mol Cell Biol *10*, 659-671.

Dornan, D., Shimizu, H., Perkins, N.D., and Hupp, T.R. (2003). DNA-dependent acetylation of p53 by the transcription coactivator p300. J Biol Chem *278*, 13431-13441.

Dorr, A., Kiermer, V., Pedal, A., Rackwitz, H.R., Henklein, P., Schubert, U., Zhou, M.M., Verdin, E., and Ott, M. (2002). Transcriptional synergy between Tat and PCAF is dependent on the binding of acetylated Tat to the PCAF bromodomain. EMBO J *21*, 2715-2723.

Du, M., and Chen, Z.J. (2018). DNA-induced liquid phase condensation of cGAS activates innate immune signaling. Science *361*, 704-709.

Elbaum-Garfinkle, S., Kim, Y., Szczepaniak, K., Chen, C.C., Eckmann, C.R., Myong, S., and Brangwynne, C.P. (2015). The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. Proc Natl Acad Sci U S A *112*, 7189-7194.

Feric, M., Vaidya, N., Harmon, T.S., Mitrea, D.M., Zhu, L., Richardson, T.M., Kriwacki, R.W., Pappu, R.V., and Brangwynne, C.P. (2016). Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell *165*, 1686-1697.

Filippakopoulos, P., and Knapp, S. (2012). The bromodomain interaction module. FEBS Lett *586*, 2692-2704.

Fischle, W., Dequiedt, F., Hendzel, M.J., Guenther, M.G., Lazar, M.A., Voelter, W., and Verdin, E. (2002). Enzymatic activity associated with class II HDACs is dependent on a multiprotein complex containing HDAC3 and SMRT/N-CoR. Mol Cell *9*, 45-57.

Floor, S.N., Condon, K.J., Sharma, D., Jankowsky, E., and Doudna, J.A. (2016). Autoinhibitory Interdomain Interactions and Subfamily-specific Extensions Redefine the Catalytic Core of the Human DEAD-box Protein DDX3. J Biol Chem *291*, 2412-2421.

Foresta, C., Ferlin, A., and Moro, E. (2000). Deletion and expression analysis of AZFa genes on the human Y chromosome revealed a major role for DBY in male infertility. Hum Mol Genet *9*, 1161-1169.

Fowler, D.M., and Fields, S. (2014). Deep mutational scanning: a new style of protein science. Nat Methods *11*,

801-807.

Franzmann, T.M., Jahnel, M., Pozniakovsky, A., Mahamid, J., Holehouse, A.S., Nuske, E., Richter, D., Baumeister, W., Grill, S.W., Pappu, R.V., *et al.* (2018). Phase separation of a yeast prion protein promotes cellular fitness. Science *359*.

Freeman Rosenzweig, E.S., Xu, B., Kuhn Cuellar, L., Martinez-Sanchez, A., Schaffer, M., Strauss, M., Cartwright, H.N., Ronceray, P., Plitzko, J.M., Forster, F., *et al.* (2017). The Eukaryotic CO2-Concentrating Organelle Is Liquid-like and Exhibits Dynamic Reorganization. Cell *171*, 148-162 e119.

Frescas, D., Valenti, L., and Accili, D. (2005). Nuclear trapping of the forkhead transcription factor FoxO1 via Sirt-dependent deacetylation promotes expression of glucogenetic genes. J Biol Chem *280*, 20589-20595.

Frey, S., Richter, R.P., and Gorlich, D. (2006). FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. Science *314*, 815-817.

Friend, L.R., Landsberg, M.J., Nouwens, A.S., Wei, Y., Rothnagel, J.A., and Smith, R. (2013). Arginine methylation of hnRNP A2 does not directly govern its subcellular localization. PLoS One *8*, e75669.

Fuhrmann, J., Clancy, K.W., and Thompson, P.R. (2015). Chemical biology of protein arginine modifications in epigenetic regulation. Chem Rev *115*, 5413-5461.

Furdas, S.D.K., S.; Sippl, W.; Jung, M. (2012). Small Molecule Inhibitors of Histone Acetyltransferases as Epigenetic Tools and Drug Candidates. Arch Pharm Chem Life Sci.

Fusco, C., Micale, L., Augello, B., Mandriani, B., Pellico, M.T., De Nittis, P., Calcagni, A., Monti, M., Cozzolino, F., Pucci, P., *et al.* (2014). HDAC6 mediates the acetylation of TRIM50. Cell Signal *26*, 363-369.

Fusco, C., Micale, L., Egorov, M., Monti, M., D'Addetta, E.V., Augello, B., Cozzolino, F., Calcagni, A., Fontana, A., Polishchuk, R.S., *et al.* (2012). The E3-ubiquitin ligase TRIM50 interacts with HDAC6 and p62, and promotes the sequestration and clearance of ubiquitinated proteins into the aggresome. PLoS One *7*, e40440.

Gao, L., Cueto, M.A., Asselbergs, F., and Atadja, P. (2002). Cloning and functional characterization of HDAC11, a novel member of the human histone deacetylase family. J Biol Chem *277*, 25748-25755.

Geissler, R., Golbik, R.P., and Behrens, S.E. (2012). The DEAD-box helicase DDX3 supports the assembly of functional 80S ribosomes. Nucleic Acids Res *40*, 4998-5011.

Geoghegan, V., Guo, A., Trudgian, D., Thomas, B., and Acuto, O. (2015). Comprehensive identification of arginine methylation in primary T cells reveals regulatory roles in cell signalling. Nat Commun *6*, 6758.

Gong, J., Sun, Z., and Li, P. (2009). CIDE proteins and metabolic disorders. Curr Opin Lipidol *20*, 121-126.

Greer, E.L., and Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. Nat Rev Genet *13*, 343-357.

Gronroos, E., Hellman, U., Heldin, C.H., and Ericsson, J. (2002). Control of Smad7 stability by competition between acetylation and ubiquitination. Mol Cell *10*, 483-493.

Grozinger, C.M., Hassig, C.A., and Schreiber, S.L. (1999). Three proteins define a class of human histone deacetylases related to yeast Hda1p. Proc Natl Acad Sci U S A *96*, 4868-4873.

Gu, L., Fullam, A., Brennan, R., and Schroder, M. (2013). Human DEAD box helicase 3 couples IkappaB kinase epsilon to interferon regulatory factor 3 activation. Mol Cell Biol *33*, 2004-2015.

Gu, W., and Roeder, R.G. (1997). Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. Cell *90*, 595-606.

Guarente, L. (2013). Calorie restriction and sirtuins revisited. Genes Dev *27*, 2072-2085.

Guenther, U.P., Weinberg, D.E., Zubradt, M.M., Tedeschi, F.A., Stawicki, B.N., Zagore, L.L., Brar, G.A., Licatalosi, D.D., Bartel, D.P., Weissman, J.S.*, et al.* (2018). The helicase Ded1p controls use of near-cognate translation initiation codons in 5' UTRs. Nature *559*, 130-134.

Gueroussov, S., Weatheritt, R.J., O'Hanlon, D., Lin, Z.Y., Narula, A., Gingras, A.C., and Blencowe, B.J. (2017). Regulatory Expansion in Mammals of Multivalent hnRNP Assemblies that Globally Control Alternative Splicing. Cell *170*, 324-339 e323.

Guo, L., Kim, H.J., Wang, H., Monaghan, J., Freyermuth, F., Sung, J.C., O'Donovan, K., Fare, C.M., Diaz, Z., Singh, N.*, et al.* (2018a). Nuclear-Import Receptors Reverse Aberrant Phase Transitions of RNA-Binding Proteins with Prion-like Domains. Cell *173*, 677-692 e620.

Guo, Q., Lehmer, C., Martinez-Sanchez, A., Rudack, T., Beck, F., Hartmann, H., Perez-Berlanga, M., Frottin, F., Hipp, M.S., Hartl, F.U.*, et al.* (2018b). In Situ Structure of Neuronal C9orf72 Poly-GA Aggregates Reveals Proteasome Recruitment. Cell *172*, 696-705 e612.

Gustafson, E.A., and Wessel, G.M. (2010). DEAD-box helicases: posttranslational regulation and function. Biochem Biophys Res Commun *395*, 1-6.

Hai, Y., and Christianson, D.W. (2016). Histone deacetylase 6 structure and molecular basis of catalysis and inhibition. Nat Chem Biol *12*, 741-747.

Hai, Y., Shinsky, S.A., Porter, N.J., and Christianson, D.W. (2017). Histone deacetylase 10 structure and molecular function as a polyamine deacetylase. Nat Commun *8*, 15368.

Hao, R., Nanduri, P., Rao, Y., Panichelli, R.S., Ito, A., Yoshida, M., and Yao, T.P. (2013). Proteasomes activate aggresome disassembly and clearance by producing unanchored ubiquitin chains. Mol Cell *51*, 819-828.

Herhaus, L., and Dikic, I. (2018). Ubiquitin-induced phase separation of p62/SQSTM1. Cell Res *28*, 389-390.

Hernandez-Vega, A., Braun, M., Scharrel, L., Jahnel, M., Wegmann, S., Hyman, B.T., Alberti, S., Diez, S., and Hyman, A.A. (2017). Local Nucleation of Microtubule Bundles through Tubulin Concentration into a Condensed Tau Phase. Cell Rep *20*, 2304-2312.

Hilliker, A., Gao, Z., Jankowsky, E., and Parker, R. (2011). The DEAD-box protein Ded1 modulates translation by the formation and resolution of an eIF4F-mRNA complex. Mol Cell *43*, 962-972.

Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., and Sharp, P.A. (2017). A Phase Separation Model for Transcriptional Control. Cell *169*, 13-23.

Hofweber, M., Hutten, S., Bourgeois, B., Spreitzer, E., Niedner-Boblenz, A., Schifferer, M., Ruepp, M.D., Simons, M., Niessing, D., Madl, T.*, et al.* (2018). Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. Cell *173*, 706-719 e713.

Hong, L., Schroth, G.P., Matthews, H.R., Yau, P., and Bradbury, E.M. (1993). Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. J Biol Chem *268*, 305-314.

Houtkooper, R.H., Pirinen, E., and Auwerx, J. (2012). Sirtuins as regulators of metabolism and healthspan. Nat Rev Mol Cell Biol *13*, 225-238.

Huang, J.S., Chao, C.C., Su, T.L., Yeh, S.H., Chen, D.S., Chen, C.T., Chen, P.J., and Jou, Y.S. (2004). Diverse cellular transformation capability of overexpressed genes in human hepatocellular carcinoma. Biochem Biophys Res Commun *315*, 950-958.

Hubbert, C., Guardiola, A., Shao, R., Kawaguchi, Y., Ito, A., Nixon, A., Yoshida, M., Wang, X.F., and Yao, T.P. (2002). HDAC6 is a microtubule-associated deacetylase. Nature *417*, 455-458.

Hubner, N.C., Bird, A.W., Cox, J., Splettstoesser, B., Bandilla, P., Poser, I., Hyman, A., and Mann, M. (2010). Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. J Cell Biol *189*, 739-754.

Hughes, M.P., Sawaya, M.R., Boyer, D.R., Goldschmidt, L., Rodriguez, J.A., Cascio, D., Chong, L., Gonen, T., and Eisenberg, D.S. (2018). Atomic structures of low-complexity protein segments reveal kinked beta sheets that assemble networks. Science *359*, 698-701.

Huo, L., Li, D., Sun, X., Shi, X., Karna, P., Yang, W., Liu, M., Qiao, W., Aneja, R., and Zhou, J. (2011). Regulation of Tat acetylation and transactivation activity by the microtubule-associated deacetylase HDAC6. J Biol Chem *286*, 9280-9286.

Hyman, A.A., Weber, C.A., and Julicher, F. (2014). Liquid-liquid phase separation in biology. Annu Rev Cell Dev Biol *30*, 39-58.

Iaconelli, J., Lalonde, J., Watmuff, B., Liu, B., Mazitschek, R., Haggarty, S.J., and Karmacharya, R. (2017). Lysine Deacetylation by HDAC6 Regulates the Kinase Activity of AKT in Human Neural Progenitor Cells. ACS Chem Biol *12*, 2139-2148.

Iborra, F.J., Pombo, A., Jackson, D.A., and Cook, P.R. (1996). Active RNA polymerases are localized within discrete transcription "factories' in human nuclei. J Cell Sci *109 (Pt 6)*, 1427-1436.

Inuzuka, H., Gao, D., Finley, L.W., Yang, W., Wan, L., Fukushima, H., Chin, Y.R., Zhai, B., Shaik, S., Lau, A.W.*, et al.* (2012). Acetylation-dependent regulation of Skp2 function. Cell *150*, 179-193.

Ishaq, M., Hu, J., Wu, X., Fu, Q., Yang, Y., Liu, Q., and Guo, D. (2008). Knockdown of cellular RNA helicase DDX3 by short hairpin RNAs suppresses HIV-1 viral replication without inducing apoptosis. Mol Biotechnol *39*, 231-238.

Ivanov, G.S., Ivanova, T., Kurash, J., Ivanov, A., Chuikov, S., Gizatullin, F., Herrera-Medina, E.M., Rauscher, F., 3rd, Reinberg, D., and Barlev, N.A. (2007). Methylation-acetylation interplay activates p53 in response to DNA damage. Mol Cell Biol *27*, 6756-6769.

Jain, A., and Vale, R.D. (2017). RNA phase transitions in repeat expansion disorders. Nature *546*, 243-247.

Jain, S., Wei, J., Mitrani, L.R., and Bishopric, N.H. (2012). Auto-acetylation stabilizes p300 in cardiac myocytes during acute oxidative stress, promoting STAT3 accumulation and cell survival. Breast Cancer Res Treat *135*, 103-114.

Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A., and Parker, R. (2016). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. Cell *164*, 487-498.

Janke, C., and Bulinski, J.C. (2011). Post-translational regulation of the microtubule cytoskeleton: mechanisms and functions. Nat Rev Mol Cell Biol *12*, 773-786.

Jedrusik-Bode, M., Studencka, M., Smolka, C., Baumann, T., Schmidt, H., Kampf, J., Paap, F., Martin, S., Tazi, J., Muller, K.M.*, et al.* (2013). The sirtuin SIRT6 regulates stress granule formation in C. elegans and mammals. J Cell Sci *126*, 5166-5177.

Jensen, T.J., Loo, M.A., Pind, S., Williams, D.B., Goldberg, A.L., and Riordan, J.R. (1995). Multiple proteolytic systems, including the proteasome, contribute to CFTR processing. Cell *83*, 129-135.

Jia, Y.J., Liu, Z.B., Wang, W.G., Sun, C.B., Wei, P., Yang, Y.L., You, M.J., Yu, B.H., Li, X.Q., and Zhou, X.Y. (2017). HDAC6 regulates microRNA-27b that suppresses proliferation, promotes apoptosis and target MET in diffuse large B-cell lymphoma. Leukemia.

Jiang, L., Gu, Z.H., Yan, Z.X., Zhao, X., Xie, Y.Y., Zhang, Z.G., Pan, C.M., Hu, Y., Cai, C.P., Dong, Y.*, et al.* (2015). Exome sequencing identifies somatic mutations of DDX3X in natural killer/T-cell lymphoma. Nat Genet *47*, 1061-1066.

Jing, E., Gesta, S., and Kahn, C.R. (2007). SIRT2 regulates adipocyte differentiation through FoxO1 acetylation/deacetylation. Cell Metab *6*, 105-114.

Johnson, J.O., Mandrioli, J., Benatar, M., Abramzon, Y., Van Deerlin, V.M., Trojanowski, J.Q., Gibbs, J.R., Brunetti, M., Gronka, S., Wuu, J.*, et al.* (2010). Exome sequencing reveals VCP mutations as a cause of familial ALS. Neuron *68*, 857-864.

Johnston, J.A., Ward, C.L., and Kopito, R.R. (1998). Aggresomes: a cellular response to misfolded proteins. J Cell Biol *143*, 1883-1898.

Kaczmarska, Z., Ortega, E., Goudarzi, A., Huang, H., Kim, S., Marquez, J.A., Zhao, Y., Khochbin, S., and Panne, D. (2017). Structure of p300 in complex with acyl-CoA variants. Nat Chem Biol *13*, 21-29.

Kaluza, D., Kroll, J., Gesierich, S., Yao, T.P., Boon, R.A., Hergenreider, E., Tjwa, M., Rossig, L., Seto, E., Augustin, H.G., *et al.* (2011). Class IIb HDAC6 regulates endothelial cell migration and angiogenesis by deacetylation of cortactin. EMBO J *30*, 4142-4156.

Kanai, Y., Dohmae, N., and Hirokawa, N. (2004). Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. Neuron *43*, 513-525.

Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J., *et al.* (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell *149*, 753-767.

Kawaguchi, Y., Kovacs, J.J., McLaurin, A., Vance, J.M., Ito, A., and Yao, T.P. (2003). The deacetylase HDAC6 regulates aggresome formation and cell viability in response to misfolded protein stress. Cell *115*, 727-738.

Kawasaki, T., and Kawai, T. (2014). Toll-like receptor signaling pathways. Front Immunol *5*, 461.

Kim, S.C., Sprung, R., Chen, Y., Xu, Y., Ball, H., Pei, J., Cheng, T., Kho, Y., Xiao, H., Xiao, L., *et al.* (2006). Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. Mol Cell *23*, 607-618.

Kim, Y., and Myong, S. (2016). RNA Remodeling Activity of DEAD Box Proteins Tuned by Protein Concentration, RNA Length, and ATP. Mol Cell *63*, 865-876.

Kirkin, V., McEwan, D.G., Novak, I., and Dikic, I. (2009). A role for ubiquitin in selective autophagy. Mol Cell *34*, 259-269.

Kopito, R.R. (2000). Aggresomes, inclusion bodies and protein aggregation. Trends Cell Biol *10*, 524-530.

Kovacs, J.J., Murphy, P.J., Gaillard, S., Zhao, X., Wu, J.T., Nicchitta, C.V., Yoshida, M., Toft, D.O., Pratt, W.B., and Yao, T.P. (2005). HDAC6 regulates Hsp90 acetylation and chaperone-dependent activation of glucocorticoid receptor. Mol Cell *18*, 601-607.

Kwiatkowski, T.J., Jr., Bosco, D.A., Leclerc, A.L., Tamrazian, E., Vanderburg, C.R., Russ, C., Davis, A., Gilchrist, J., Kasarskis, E.J., Munsat, T., *et al.* (2009). Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. Science *323*, 1205-1208.

Kwon, I., Kato, M., Xiang, S., Wu, L., Theodoropoulos, P., Mirzaei, H., Han, T., Xie, S., Corden, J.L., and McKnight, S.L. (2013). Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. Cell *155*, 1049-1060.

Kwon, I., Xiang, S., Kato, M., Wu, L., Theodoropoulos, P., Wang, T., Kim, J., Yun, J., Xie, Y., and McKnight, S.L. (2014). Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. Science *345*, 1139-1145.

Kwon, S., Zhang, Y., and Matthias, P. (2007). The deacetylase HDAC6 is a novel critical component of stress granules involved in the stress response. Genes Dev *21*, 3381-3394.

Labokha, A.A., Gradmann, S., Frey, S., Hulsmann, B.B., Urlaub, H., Baldus, M., and Gorlich, D. (2013). Systematic analysis of barrier-forming FG hydrogels from Xenopus nuclear pore complexes. EMBO J *32*, 204-218.

Lagier-Tourenne, C., and Cleveland, D.W. (2009). Rethinking ALS: the FUS about TDP-43. Cell *136*, 1001-1004.

Lai, M.C., Chang, W.C., Shieh, S.Y., and Tarn, W.Y. (2010). DDX3 regulates cell growth through translational control of cyclin E1. Mol Cell Biol *30*, 5444-5453.

Lai, M.C., Lee, Y.H., and Tarn, W.Y. (2008). The DEAD-box RNA helicase DDX3 associates with export messenger ribonucleoproteins as well as tip-associated protein and participates in translational control. Mol Biol Cell *19*, 3847-3858.

Langdon, E.M., Qiu, Y., Ghanbari Niaki, A., McLaughlin, G.A., Weidmann, C.A., Gerbich, T.M., Smith, J.A., Crutchley, J.M., Termini, C.M., Weeks, K.M.*, et al.* (2018). mRNA structure determines specificity of a polyQ-driven phase separation. Science *360*, 922-927.

Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. Nature *547*, 236-240.

Le Hir, H., and Andersen, G.R. (2008). Structural insights into the exon junction complex. Curr Opin Struct Biol *18*, 112-119.

Lee, C.S., Dias, A.P., Jedrychowski, M., Patel, A.H., Hsu, J.L., and Reed, R. (2008). Human DDX3 functions in translation and interacts with the translation initiation factor eIF3. Nucleic Acids Res *36*, 4708-4718.

Lee, J.Y., Kapur, M., Li, M., Choi, M.C., Choi, S., Kim, H.J., Kim, I., Lee, E., Taylor, J.P., and Yao, T.P. (2014a). MFN1 deacetylation activates adaptive mitochondrial fusion and protects metabolically challenged mitochondria. J Cell Sci *127*, 4954-4963.

Lee, J.Y., Koga, H., Kawaguchi, Y., Tang, W., Wong, E., Gao, Y.S., Pandey, U.B., Kaushik, S., Tresse, E., Lu, J.*, et al.* (2010). HDAC6 controls autophagosome maturation essential for ubiquitin-selective quality-control autophagy. EMBO J *29*, 969-980.

Lee, K.H., Zhang, P., Kim, H.J., Mitrea, D.M., Sarkar, M., Freibaum, B.D., Cika, J., Coughlin, M., Messing, J., Molliex, A.*, et al.* (2016). C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. Cell *167*, 774-788 e717.

Lee, Y.J., Hsieh, W.Y., Chen, L.Y., and Li, C. (2012). Protein arginine methylation of SERBP1 by protein arginine methyltransferase 1 affects cytoplasmic/nuclear distribution. J Cell Biochem *113*, 2721-2728.

Lee, Y.J., Wei, H.M., Chen, L.Y., and Li, C. (2014b). Localization of SERBP1 in stress granules and nucleoli. FEBS J *281*, 352-364.

Leung, D.W., and Amarasinghe, G.K. (2012). Structural insights into RNA recognition and activation of RIG-I-like receptors. Curr Opin Struct Biol *22*, 297-303.

Levine, M., Cattoglio, C., and Tjian, R. (2014). Looping back to leap forward: transcription enters a new era. Cell *157*, 13-25.

Li, G.M. (2008). Mechanisms and functions of DNA mismatch repair. Cell Res *18*, 85-98.

Li, L., Fang, R., Liu, B., Shi, H., Wang, Y., Zhang, W., Zhang, X., and Ye, L. (2016a). Deacetylation of tumor-suppressor MST1 in Hippo pathway induces its degradation through HBXIP-elevated HDAC6 in promotion of breast cancer growth. Oncogene *35*, 4048-4057.

Li, P., Banjade, S., Cheng, H.C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J.V., King, D.S., Banani, S.F.*, et al.* (2012). Phase transitions in the assembly of multivalent signalling proteins. Nature *483*, 336-340.

Li, P., Kurata, Y., Endang, M., Ninomiya, H., Higaki, K., Taufiq, F., Morikawa, K., Shirayoshi, Y., Horie, M., and Hisatome, I. (2018). Restoration of mutant hERG stability by inhibition of HDAC6. J Mol Cell Cardiol.

Li, S., Yang, P., Tian, E., and Zhang, H. (2013). Arginine methylation modulates autophagic degradation of PGL granules in C. elegans. Mol Cell *52*, 421-433.

Li, X., Tao, Y., Murphy, J.W., Scherer, A.N., Lam, T.T., Marshall, A.G., Koleske, A.J., and Boggon, T.J. (2017). The repeat region of cortactin is intrinsically disordered in solution. Sci Rep *7*, 16696.

Li, Y., Sabari, B.R., Panchenko, T., Wen, H., Zhao, D., Guan, H., Wan, L., Huang, H., Tang, Z., Zhao, Y., *et al.* (2016b). Molecular Coupling of Histone Crotonylation and Active Transcription by AF9 YEATS Domain. Mol Cell *62*, 181-193.

Li, Y., Wen, H., Xi, Y., Tanaka, K., Wang, H., Peng, D., Ren, Y., Jin, Q., Dent, S.Y., Li, W., *et al.* (2014). AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation. Cell *159*, 558-571.

Li, Y., Zhang, X., Polakiewicz, R.D., Yao, T.P., and Comb, M.J. (2008). HDAC6 is required for epidermal growth factor-induced beta-catenin nuclear localization. J Biol Chem *283*, 12686-12690.

Li, Z.Y., Li, Q.Z., Chen, L., Chen, B.D., Zhang, C., Wang, X., and Li, W.P. (2016c). HPOB, an HDAC6 inhibitor, attenuates corticosterone-induced injury in rat adrenal pheochromocytoma PC12 cells by inhibiting mitochondrial GR translocation and the intrinsic apoptosis pathway. Neurochem Int *99*, 239-251.

Lin, Y., Mori, E., Kato, M., Xiang, S., Wu, L., Kwon, I., and McKnight, S.L. (2016). Toxic PR Poly-Dipeptides Encoded by the C9orf72 Repeat Expansion Target LC Domain Polymers. Cell *167*, 789-802 e712.

Lin, Y., Protter, D.S., Rosen, M.K., and Parker, R. (2015). Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Mol Cell *60*, 208-219.

Linder, P. (2006). Dead-box proteins: a family affair--active and passive players in RNP-remodeling. Nucleic Acids Res *34*, 4168-4180.

Linder, P., and Jankowsky, E. (2011). From unwinding to clamping - the DEAD box RNA helicase family. Nat Rev Mol Cell Biol *12*, 505-516.

Liu, F., You, X., Wang, Y., Liu, Q., Liu, Y., Zhang, S., Chen, L., Zhang, X., and Ye, L. (2014). The oncoprotein HBXIP enhances angiogenesis and growth of breast cancer through modulating FGF8 and VEGF. Carcinogenesis *35*, 1144-1153.

Liu, N., Xiong, Y., Li, S., Ren, Y., He, Q., Gao, S., Zhou, J., and Shui, W. (2015). New HDAC6-mediated deacetylation sites of tubulin in the mouse brain identified by quantitative mass spectrometry. Sci Rep *5*, 16869.

Lu, H., Yu, D., Hansen, A.S., Ganguly, S., Liu, R., Heckert, A., Darzacq, X., and Zhou, Q. (2018). Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. Nature *558*, 318-323.

Lu, K.P., and Zhou, X.Z. (2007). The prolyl isomerase PIN1: a pivotal new twist in phosphorylation signalling and disease. Nat Rev Mol Cell Biol *8*, 904-916.

Lu, Z., Ghosh, S., Wang, Z., and Hunter, T. (2003). Downregulation of caveolin-1 function by EGF leads to the loss of E-cadherin, increased transcriptional activity of beta-catenin, and enhanced tumor cell invasion. Cancer Cell *4*, 499-515.

Mackenzie, I.R., Nicholson, A.M., Sarkar, M., Messing, J., Purice, M.D., Pottier, C., Annu, K., Baker, M., Perkerson, R.B., Kurti, A.*, et al.* (2017). TIA1 Mutations in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Promote Phase Separation and Alter Stress Granule Dynamics. Neuron *95*, 808-816 e809.

Maharana, S., Wang, J., Papadopoulos, D.K., Richter, D., Pozniakovsky, A., Poser, I., Bickle, M., Rizk, S., Guillen-Boixet, J., Franzmann, T.M.*, et al.* (2018). RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science *360*, 918-921.

Markmiller, S., Soltanieh, S., Server, K.L., Mak, R., Jin, W., Fang, M.Y., Luo, E.C., Krach, F., Yang, D., Sen, A.*, et al.* (2018). Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. Cell *172*, 590-604 e513.

Maruta, H., Greer, K., and Rosenbaum, J.L. (1986). The acetylation of alpha-tubulin and its relationship to the assembly and disassembly of microtubules. J Cell Biol *103*, 571-579.

Marzahn, M.R., Marada, S., Lee, J., Nourse, A., Kenrick, S., Zhao, H., Ben-Nissan, G., Kolaitis, R.M., Peters, J.L., Pounds, S.*, et al.* (2016). Higher-order oligomerization promotes localization of SPOP to liquid nuclear speckles. EMBO J *35*, 1254-1275.

Matsuzaki, H., Daitoku, H., Hatta, M., Aoyama, H., Yoshimochi, K., and Fukamizu, A. (2005). Acetylation of Foxo1 alters its DNA-binding ability and sensitivity to phosphorylation. Proc Natl Acad Sci U S A *102*, 11278-11283.

Matthias, P., Yoshida, M., and Khochbin, S. (2008). HDAC6 a new cellular stress surveillance factor. Cell Cycle *7*, 7-10.

Mazroui, R., Di Marco, S., Kaufman, R.J., and Gallouzi, I.E. (2007). Inhibition of the ubiquitin-proteasome system induces stress granule formation. Mol Biol Cell *18*, 2603-2618.

McGurk, L., Gomes, E., Guo, L., Mojsilovic-Petrovic, J., Tran, V., Kalb, R.G., Shorter, J., and Bonini, N.M. (2018). Poly(ADP-Ribose) Prevents Pathological Phase Separation of TDP-43 by Promoting Liquid Demixing and Stress Granule Localization. Mol Cell *71*, 703-717 e709.

McKinsey, T.A., Zhang, C.L., Lu, J., and Olson, E.N. (2000). Signal-dependent nuclear export of a histone

deacetylase regulates muscle differentiation. Nature *408*, 106-111.

Medler, T.R., Craig, J.M., Fiorillo, A.A., Feeney, Y.B., Harrell, J.C., and Clevenger, C.V. (2016). HDAC6 Deacetylates HMGN2 to Regulate Stat5a Activity and Breast Cancer Growth. Mol Cancer Res *14*, 994-1008.

Meng, Z., Jia, L.F., and Gan, Y.H. (2016). PTEN activation through K163 acetylation by inhibiting HDAC6 contributes to tumour inhibition. Oncogene *35*, 2333-2344.

Merz, C., Urlaub, H., Will, C.L., and Luhrmann, R. (2007). Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. RNA *13*, 116-128.

Messaoudi, K., Ali, A., Ishaq, R., Palazzo, A., Sliwa, D., Bluteau, O., Souquere, S., Muller, D., Diop, K.M., Rameau, P.*, et al.* (2017). Critical role of the HDAC6-cortactin axis in human megakaryocyte maturation leading to a proplatelet-formation defect. Nat Commun *8*, 1786.

Milovanovic, D., and De Camilli, P. (2017). Synaptic Vesicle Clusters at Synapses: A Distinct Liquid Phase? Neuron *93*, 995-1002.

Milovanovic, D., Wu, Y., Bian, X., and De Camilli, P. (2018). A liquid phase of synapsin and lipid vesicles. Science *361*, 604-607.

Min, S.W., Cho, S.H., Zhou, Y., Schroeder, S., Haroutunian, V., Seeley, W.W., Huang, E.J., Shen, Y., Masliah, E., Mukherjee, C.*, et al.* (2010). Acetylation of tau inhibits its degradation and contributes to tauopathy. Neuron *67*, 953-966.

Mitchell, P.J., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science *245*, 371-378.

Mittasch, M., Gross, P., Nestler, M., Fritsch, A.W., Iserman, C., Kar, M., Munder, M., Voigt, A., Alberti, S., Grill, S.W.*, et al.* (2018). Non-invasive perturbations of intracellular flow reveal physical principles of cell organization. Nat Cell Biol *20*, 344-351.

Miyake, Y., Keusch, J.J., Wang, L., Saito, M., Hess, D., Wang, X., Melancon, B.J., Helquist, P., Gut, H., and Matthias, P. (2016). Structural insights into HDAC6 tubulin deacetylation and its selective inhibition. Nat Chem Biol *12*, 748-754.

Mizushima, N., and Komatsu, M. (2011). Autophagy: renovation of cells and tissues. Cell *147*, 728-741.

Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T., and Taylor, J.P. (2015). Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. Cell *163*, 123-133.

Monahan, Z., Ryan, V.H., Janke, A.M., Burke, K.A., Rhoads, S.N., Zerze, G.H., O'Meally, R., Dignon, G.L., Conicella, A.E., Zheng, W.*, et al.* (2017). Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. EMBO J *36*, 2951-2967.

Mooney, S.M., Goel, A., D'Assoro, A.B., Salisbury, J.L., and Janknecht, R. (2010). Pleiotropic effects of p300-mediated acetylation on p68 and p72 RNA helicase. J Biol Chem *285*, 30443-30452.

Moremen, K.W., Tiemeyer, M., and Nairn, A.V. (2012). Vertebrate protein glycosylation: diversity, synthesis and function. Nat Rev Mol Cell Biol *13*, 448-462.

Morrison, D.K. (2009). The 14-3-3 proteins: integrators of diverse signaling cues that impact cell fate and cancer development. Trends Cell Biol *19*, 16-23.

Mortenson, J.B., Heppler, L.N., Banks, C.J., Weerasekara, V.K., Whited, M.D., Piccolo, S.R., Johnson, W.E., Thompson, J.W., and Andersen, J.L. (2015). Histone deacetylase 6 (HDAC6) promotes the pro-survival activity of 14-3-3zeta via deacetylation of lysines within the 14-3-3zeta binding pocket. J Biol Chem *290*, 12487-12496.

Mugler, C.F., Hondele, M., Heinrich, S., Sachdev, R., Vallotton, P., Koek, A.Y., Chan, L.Y., and Weis, K. (2016). ATPase activity of the DEAD-box protein Dhh1 controls processing body formation. Elife *5*.

Mujtaba, S., He, Y., Zeng, L., Farooq, A., Carlson, J.E., Ott, M., Verdin, E., and Zhou, M.M. (2002). Structural basis of lysine-acetylated HIV-1 Tat recognition by PCAF bromodomain. Mol Cell *9*, 575-586.

Mujtaba, S., He, Y., Zeng, L., Yan, S., Plotnikova, O., Sachchidanand, Sanchez, R., Zeleznik-Le, N.J., Ronai, Z., and Zhou, M.M. (2004). Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation. Mol Cell *13*, 251-263.

Murakami, T., Qamar, S., Lin, J.Q., Schierle, G.S., Rees, E., Miyashita, A., Costa, A.R., Dodd, R.B., Chan, F.T., Michel, C.H.*, et al.* (2015). ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. Neuron *88*, 678-690.

Murray, D.T., Kato, M., Lin, Y., Thurber, K.R., Hung, I., McKnight, S.L., and Tycko, R. (2017). Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. Cell *171*, 615-627 e616.

Neumann, H., Peak-Chew, S.Y., and Chin, J.W. (2008). Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. Nat Chem Biol *4*, 232-234.

Neumann, M., Sampathu, D.M., Kwong, L.K., Truax, A.C., Micsenyi, M.C., Chou, T.T., Bruce, J., Schuck, T., Grossman, M., Clark, C.M.*, et al.* (2006). Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science *314*, 130-133.

Nishida, Y., Rardin, M.J., Carrico, C., He, W., Sahu, A.K., Gut, P., Najjar, R., Fitch, M., Hellerstein, M., Gibson, B.W.*, et al.* (2015). SIRT5 Regulates both Cytosolic and Mitochondrial Protein Malonylation with Glycolysis as a Major Target. Mol Cell *59*, 321-332.

Noack, M., Leyk, J., and Richter-Landsberg, C. (2014). HDAC6 inhibition results in tau acetylation and modulates tau phosphorylation and degradation in oligodendrocytes. Glia *62*, 535-547.

Nogues, L., Reglero, C., Rivas, V., Salcedo, A., Lafarga, V., Neves, M., Ramos, P., Mendiola, M., Berjon, A., Stamatakis, K.*, et al.* (2016). G Protein-coupled Receptor Kinase 2 (GRK2) Promotes Breast Tumorigenesis Through a HDAC6-Pin1 Axis. EBioMedicine *13*, 132-145.

Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D.*, et al.* (2015). Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. Mol Cell *57*, 936-947.

Ohn, T., Kedersha, N., Hickman, T., Tisdale, S., and Anderson, P. (2008). A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly. Nat Cell Biol *10*, 1224-1231.

Oshiumi, H., Sakai, K., Matsumoto, M., and Seya, T. (2010). DEAD/H BOX 3 (DDX3) helicase binds the RIG-I adaptor IPS-1 to up-regulate IFN-beta-inducing potential. Eur J Immunol *40*, 940-948.

Ostapcuk, V., Mohn, F., Carl, S.H., Basters, A., Hess, D., Iesmantavicius, V., Lampersberger, L., Flemr, M., Pandey, A., Thoma, N.H.*, et al.* (2018). Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. Nature *557*, 739-743.

Ouyang, H., Ali, Y.O., Ravichandran, M., Dong, A., Qiu, W., MacKenzie, F., Dhe-Paganon, S., Arrowsmith, C.H., and Zhai, R.G. (2012). Protein aggregates are recruited to aggresome by histone deacetylase 6 via unanchored ubiquitin C termini. J Biol Chem *287*, 2317-2327.

Owsianka, A.M., and Patel, A.H. (1999). Hepatitis C virus core protein interacts with a human DEAD box

protein DDX3. Virology *257*, 330-340.

Page, L.J., Sowerby, P.J., Lui, W.W., and Robinson, M.S. (1999). Gamma-synergin: an EH domain-containing protein that interacts with gamma-adaptin. J Cell Biol *146*, 993-1004.

Papantonis, A., and Cook, P.R. (2013). Transcription factories: genome organization and gene regulation. Chem Rev *113*, 8683-8705.

Park, J., Chen, Y., Tishkoff, D.X., Peng, C., Tan, M., Dai, L., Xie, Z., Zhang, Y., Zwaans, B.M., Skinner, M.E.*, et al.* (2013). SIRT5-mediated lysine desuccinylation impacts diverse metabolic pathways. Mol Cell *50*, 919-930.

Park, S.H., Lee, S.G., Kim, Y., and Song, K. (1998). Assignment of a human putative RNA helicase gene, DDX3, to human X chromosome bands p11.3-->p11.23. Cytogenet Cell Genet *81*, 178-179.

Parmigiani, R.B., Xu, W.S., Venta-Perez, G., Erdjument-Bromage, H., Yaneva, M., Tempst, P., and Marks, P.A. (2008). HDAC6 is a specific deacetylase of peroxiredoxins and is involved in redox regulation. Proc Natl Acad Sci U S A *105*, 9633-9638.

Parthun, M.R. (2007). Hat1: the emerging cellular roles of a type B histone acetyltransferase. Oncogene *26*, 5319-5328.

Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T.M.*, et al.* (2015). A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell *162*, 1066-1077.

Pause, A., and Sonenberg, N. (1992). Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. EMBO J *11*, 2643-2654.

Pawson, T., and Scott, J.D. (2005). Protein phosphorylation in signaling--50 years and counting. Trends Biochem Sci *30*, 286-290.

Peskett, T.R., Rau, F., O'Driscoll, J., Patani, R., Lowe, A.R., and Saibil, H.R. (2018). A Liquid to Solid Phase Transition Underlying Pathological Huntingtin Exon1 Aggregation. Mol Cell *70*, 588-601 e586.

Platta, H.W., El Magraoui, F., Schlee, D., Grunau, S., Girzalsky, W., and Erdmann, R. (2007). Ubiquitination of the peroxisomal import receptor Pex5p is required for its recycling. J Cell Biol *177*, 197-204.

Portran, D., Schaedel, L., Xu, Z., Thery, M., and Nachury, M.V. (2017). Tubulin acetylation protects long-lived microtubules against mechanical ageing. Nat Cell Biol *19*, 391-398.

Protter, D.S.W., and Parker, R. (2016). Principles and Properties of Stress Granules. Trends Cell Biol *26*, 668-679.

Pugh, T.J., Weeraratne, S.D., Archer, T.C., Pomeranz Krummel, D.A., Auclair, D., Bochicchio, J., Carneiro, M.O., Carter, S.L., Cibulskis, K., Erlich, R.L.*, et al.* (2012). Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. Nature *488*, 106-110.

Pylayeva-Gupta, Y., Grabocka, E., and Bar-Sagi, D. (2011). RAS oncogenes: weaving a tumorigenic web. Nat Rev Cancer *11*, 761-774.

Qamar, S., Wang, G., Randle, S.J., Ruggeri, F.S., Varela, J.A., Lin, J.Q., Phillips, E.C., Miyashita, A., Williams, D., Strohl, F.*, et al.* (2018). FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-pi Interactions. Cell *173*, 720-734 e715.

Qian, H., Chen, Y., Nian, Z., Su, L., Yu, H., Chen, F.J., Zhang, X., Xu, W., Zhou, L., Liu, J.*, et al.* (2017). HDAC6-mediated acetylation of lipid droplet-binding protein CIDEC regulates fat-induced lipid storage. J Clin Invest *127*, 1353-1369.

Rai, A.K., Chen, J.X., Selbach, M., and Pelkmans, L. (2018). Kinase-controlled phase transition of membraneless organelles in mitosis. Nature *559*, 211-216.

Ramaswami, M., Taylor, J.P., and Parker, R. (2013). Altered ribostasis: RNA-protein granules in degenerative disorders. Cell *154*, 727-736.

Rao, R., Nalluri, S., Kolhe, R., Yang, Y., Fiskus, W., Chen, J., Ha, K., Buckley, K.M., Balusu, R., Coothankandaswamy, V.*, et al.* (2010). Treatment with panobinostat induces glucose-regulated protein 78 acetylation and endoplasmic reticulum stress in breast cancer cells. Mol Cancer Ther *9*, 942-952.

Ravarani, C.N., Erkina, T.Y., De Baets, G., Dudman, D.C., Erkine, A.M., and Babu, M.M. (2018). High-throughput discovery of functional disordered regions: investigation of transactivation domains. Mol Syst Biol *14*, e8190.

Reichert, N., Choukrallah, M.A., and Matthias, P. (2012). Multiple roles of class I HDACs in proliferation, differentiation, and development. Cell Mol Life Sci *69*, 2173-2187.

Riback, J.A., Katanski, C.D., Kear-Scott, J.L., Pilipenko, E.V., Rojek, A.E., Sosnick, T.R., and Drummond, D.A. (2017). Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. Cell *168*, 1028-1040 e1019.

Riolo, M.T., Cooper, Z.A., Holloway, M.P., Cheng, Y., Bianchi, C., Yakirevich, E., Ma, L., Chin, Y.E., and Altura, R.A. (2012). Histone deacetylase 6 (HDAC6) deacetylates survivin for its nuclear export in breast cancer. J Biol Chem *287*, 10885-10893.

Robinson, G., Parker, M., Kranenburg, T.A., Lu, C., Chen, X., Ding, L., Phoenix, T.N., Hedlund, E., Wei, L., Zhu, X.*, et al.* (2012). Novel mutations target distinct subgroups of medulloblastoma. Nature *488*, 43-48.

Rollins, M., Huard, S., Morettin, A., Takuski, J., Pham, T.T., Fullerton, M.D., Cote, J., and Baetz, K. (2017). Lysine acetyltransferase NuA4 and acetyl-CoA regulate glucose-deprived stress granule formation in Saccharomyces cerevisiae. PLoS Genet *13*, e1006626.

Ron, D., and Walter, P. (2007). Signal integration in the endoplasmic reticulum unfolded protein response. Nat Rev Mol Cell Biol *8*, 519-529.

Ryan, V.H., Dignon, G.L., Zerze, G.H., Chabata, C.V., Silva, R., Conicella, A.E., Amaya, J., Burke, K.A., Mittal, J., and Fawzi, N.L. (2018). Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. Mol Cell *69*, 465-479 e467.

Ryu, H.W., Shin, D.H., Lee, D.H., Choi, J., Han, G., Lee, K.Y., and Kwon, S.H. (2017a). HDAC6 deacetylates p53 at lysines 381/382 and differentially coordinates p53-induced apoptosis. Cancer Lett *391*, 162-171.

Ryu, H.W., Shin, D.H., Lee, D.H., Won, H.R., and Kwon, S.H. (2018). A potent hydroxamic acid-based, small-molecule inhibitor A452 preferentially inhibits HDAC6 activity and induces cytotoxicity toward cancer cells irrespective of p53 status. Carcinogenesis *39*, 72-83.

Ryu, H.W., Won, H.R., Lee, D.H., and Kwon, S.H. (2017b). HDAC6 regulates sensitivity to cell death in response to stress and post-stress recovery. Cell Stress Chaperones *22*, 253-261.

Sabari, B.R., Dall'Agnese, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C.*, et al.* (2018). Coactivator condensation at super-enhancers links phase separation and gene control. Science *361*.

Sabari, B.R., Tang, Z., Huang, H., Yong-Gonzalez, V., Molina, H., Kong, H.E., Dai, L., Shimada, M., Cross, J.R., Zhao, Y.*, et al.* (2015). Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone

crotonylation. Mol Cell *58*, 203-215.

Sabari, B.R., Zhang, D., Allis, C.D., and Zhao, Y. (2017). Metabolic regulation of gene expression through histone acylations. Nat Rev Mol Cell Biol *18*, 90-101.

Saha, S., Weber, C.A., Nousch, M., Adame-Arana, O., Hoege, C., Hein, M.Y., Osborne-Nishimura, E., Mahamid, J., Jahnel, M., Jawerth, L.*, et al.* (2016). Polar Positioning of Phase-Separated Liquid Compartments in Cells Regulated by an mRNA Competition Mechanism. Cell *166*, 1572-1584 e1516.

Saito, M., Hess, D., Eglinger, J., Fritsch, A.W., Kreysing, M., Weinert, B.T., Choudhary, C., and Matthias, P. (2019). Acetylation of intrinsically disordered regions regulates phase separation. Nat Chem Biol *15*, 51-61.

Schmidt, H.B., and Gorlich, D. (2016). Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles. Trends Biochem Sci *41*, 46-61.

Scholz, C., Weinert, B.T., Wagner, S.A., Beli, P., Miyake, Y., Qi, J., Jensen, L.J., Streicher, W., McCarthy, A.R., Westwood, N.J.*, et al.* (2015). Acetylation site specificities of lysine deacetylase inhibitors in human cells. Nat Biotechnol *33*, 415-423.

Schroder, M., Baran, M., and Bowie, A.G. (2008). Viral targeting of DEAD box protein 3 reveals its role in TBK1/IKKepsilon-mediated IRF activation. EMBO J *27*, 2147-2157.

Schroder, S., Herker, E., Itzen, F., He, D., Thomas, S., Gilchrist, D.A., Kaehlcke, K., Cho, S., Pollard, K.S., Capra, J.A.*, et al.* (2013). Acetylation of RNA polymerase II regulates growth-factor-induced gene transcription in mammalian cells. Mol Cell *52*, 314-324.

Schwartz, J.C., Ebmeier, C.C., Podell, E.R., Heimiller, J., Taatjes, D.J., and Cech, T.R. (2012). FUS binds the CTD of RNA polymerase II and regulates its phosphorylation at Ser2. Genes Dev *26*, 2690-2695.

Schwer, B., Bunkenborg, J., Verdin, R.O., Andersen, J.S., and Verdin, E. (2006). Reversible lysine acetylation controls the activity of the mitochondrial enzyme acetyl-CoA synthetase 2. Proc Natl Acad Sci U S A *103*, 10224-10229.

Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S., and Yokoyama, S. (2006). Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. Cell *125*, 287-300.

Serizawa, H., Conaway, J.W., and Conaway, R.C. (1993). Phosphorylation of C-terminal domain of RNA polymerase II is not required in basal transcription. Nature *363*, 371-374.

Sharma, D., and Jankowsky, E. (2014). The Ded1/DDX3 subfamily of DEAD-box RNA helicases. Crit Rev Biochem Mol Biol *49*, 343-360.

Sheu-Gruttadauria, J., and MacRae, I.J. (2018). Phase Transitions in the Assembly and Function of Human miRISC. Cell *173*, 946-957 e916.

Shi, J., Wang, Y., Zeng, L., Wu, Y., Deng, J., Zhang, Q., Lin, Y., Li, J., Kang, T., Tao, M.*, et al.* (2014). Disrupting the interaction of BRD4 with diacetylated Twist suppresses tumorigenesis in basal-like breast cancer. Cancer Cell *25*, 210-225.

Shih, J.W., Tsai, T.Y., Chao, C.H., and Wu Lee, Y.H. (2008). Candidate tumor suppressor DDX3 RNA helicase specifically represses cap-dependent translation by acting as an eIF4E inhibitory protein. Oncogene *27*, 700-714.

Shih, J.W., Wang, W.T., Tsai, T.Y., Kuo, C.Y., Li, H.K., and Wu Lee, Y.H. (2012). Critical roles of RNA helicase DDX3 and its interactions with eIF4E/PABP1 in stress granule assembly and stress response. Biochem J *441*, 119-129.

Shin, Y., Berry, J., Pannucci, N., Haataja, M.P., Toettcher, J.E., and Brangwynne, C.P. (2017). Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. Cell *168*, 159-171 e114.

Shin, Y., and Brangwynne, C.P. (2017). Liquid phase condensation in cell physiology and disease. Science *357*.

Simon, J.R., Carroll, N.J., Rubinstein, M., Chilkoti, A., and Lopez, G.P. (2017). Programming molecular self-assembly of intrinsically disordered proteins containing sequences of low complexity. Nat Chem *9*, 509-515.

Simonsson, M., Heldin, C.H., Ericsson, J., and Gronroos, E. (2005). The balance between acetylation and deacetylation controls Smad7 stability. J Biol Chem *280*, 21797-21803.

Skultetyova, L., Ustinova, K., Kutil, Z., Novakova, Z., Pavlicek, J., Mikesova, J., Trapl, D., Baranova, P., Havlinova, B., Hubalek, M.*, et al.* (2017). Human histone deacetylase 6 shows strong preference for tubulin dimers over assembled microtubules. Sci Rep *7*, 11547.

Snijders Blok, L., Madsen, E., Juusola, J., Gilissen, C., Baralle, D., Reijnders, M.R., Venselaar, H., Helsmoortel, C., Cho, M.T., Hoischen, A.*, et al.* (2015). Mutations in DDX3X Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on Wnt Signaling. Am J Hum Genet *97*, 343-352.

Song, M.S., Salmena, L., and Pandolfi, P.P. (2012). The functions and regulation of the PTEN tumour suppressor. Nat Rev Mol Cell Biol *13*, 283-296.

Soto-Rifo, R., and Ohlmann, T. (2013). The role of the DEAD-box RNA helicase DDX3 in mRNA metabolism. Wiley Interdiscip Rev RNA *4*, 369-385.

Soto-Rifo, R., Rubilar, P.S., Limousin, T., de Breyne, S., Decimo, D., and Ohlmann, T. (2012). DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. EMBO J *31*, 3745-3756.

Soulat, D., Burckstummer, T., Westermayer, S., Goncalves, A., Bauch, A., Stefanovic, A., Hantschel, O., Bennett, K.L., Decker, T., and Superti-Furga, G. (2008). The DEAD-box helicase DDX3X is a critical component of the TANK-binding kinase 1-dependent innate immune response. EMBO J *27*, 2135-2146.

Staby, L., O'Shea, C., Willemoes, M., Theisen, F., Kragelund, B.B., and Skriver, K. (2017). Eukaryotic transcription factors: paradigms of protein intrinsic disorder. Biochem J *474*, 2509-2532.

Strambio-De-Castillia, C., Niepel, M., and Rout, M.P. (2010). The nuclear pore complex: bridging nuclear transport and gene regulation. Nat Rev Mol Cell Biol *11*, 490-501.

Stransky, N., Egloff, A.M., Tward, A.D., Kostic, A.D., Cibulskis, K., Sivachenko, A., Kryukov, G.V., Lawrence, M.S., Sougnez, C., McKenna, A.*, et al.* (2011). The mutational landscape of head and neck squamous cell carcinoma. Science *333*, 1157-1160.

Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., and Karpen, G.H. (2017). Phase separation drives heterochromatin domain formation. Nature *547*, 241-245.

Su, X., Ditlev, J.A., Hui, E., Xing, W., Banjade, S., Okrut, J., King, D.S., Taunton, J., Rosen, M.K., and Vale, R.D. (2016). Phase separation of signaling molecules promotes T cell receptor signal transduction. Science *352*, 595-599.

Subramanian, C., Jarzembowski, J.A., Opipari, A.W., Jr., Castle, V.P., and Kwok, R.P. (2011). HDAC6 deacetylates Ku70 and regulates Ku70-Bax binding in neuroblastoma. Neoplasia *13*, 726-734.

Sun, D., Wu, R., Zheng, J., Li, P., and Yu, L. (2018). Polyubiquitin chain-induced p62 phase separation drives autophagic cargo segregation. Cell Res *28*, 405-415.

Swatek, K.N., and Komander, D. (2016). Ubiquitin modifications. Cell Res *26*, 399-422.

Tan, M., Peng, C., Anderson, K.A., Chhoy, P., Xie, Z., Dai, L., Park, J., Chen, Y., Huang, H., Zhang, Y., *et al.* (2014). Lysine glutarylation is a protein posttranslational modification regulated by SIRT5. Cell Metab *19*, 605-617.

Thompson, P.R., Wang, D., Wang, L., Fulco, M., Pediconi, N., Zhang, D., An, W., Ge, Q., Roeder, R.G., Wong, J., *et al.* (2004). Regulation of the p300 HAT domain via a novel activation loop. Nat Struct Mol Biol *11*, 308-315.

Topisirovic, I., Siddiqui, N., Lapointe, V.L., Trost, M., Thibault, P., Bangeranye, C., Pinol-Roma, S., and Borden, K.L. (2009). Molecular dissection of the eukaryotic initiation factor 4E (eIF4E) export-competent RNP. EMBO J *28*, 1087-1098.

Tourriere, H., Chebli, K., Zekri, L., Courselaud, B., Blanchard, J.M., Bertrand, E., and Tazi, J. (2003). The RasGAP-associated endoribonuclease G3BP assembles stress granules. J Cell Biol *160*, 823-831.

Tweedie-Cullen, R.Y., Brunner, A.M., Grossmann, J., Mohanna, S., Sichau, D., Nanni, P., Panse, C., and Mansuy, I.M. (2012). Identification of combinatorial patterns of post-translational modifications on individual histones in the mouse brain. PLoS One *7*, e36980.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods *13*, 731-740.

Valentin-Vega, Y.A., Wang, Y.D., Parker, M., Patmore, D.M., Kanagaraj, A., Moore, J., Rusch, M., Finkelstein, D., Ellison, D.W., Gilbertson, R.J., *et al.* (2016). Cancer-associated DDX3X mutations drive stress granule assembly and impair global translation. Sci Rep *6*, 25996.

Vance, C., Rogelj, B., Hortobagyi, T., De Vos, K.J., Nishimura, A.L., Sreedharan, J., Hu, X., Smith, B., Ruddy, D., Wright, P., *et al.* (2009). Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science *323*, 1208-1211.

Vashist, S., Urena, L., Chaudhry, Y., and Goodfellow, I. (2012). Identification of RNA-protein interaction networks involved in the norovirus life cycle. J Virol *86*, 11977-11990.

Vega, R.B., Harrison, B.C., Meadows, E., Roberts, C.R., Papst, P.J., Olson, E.N., and McKinsey, T.A. (2004a). Protein kinases C and D mediate agonist-dependent cardiac hypertrophy through nuclear export of histone deacetylase 5. Mol Cell Biol *24*, 8374-8385.

Vega, R.B., Matsuda, K., Oh, J., Barbosa, A.C., Yang, X., Meadows, E., McAnally, J., Pomajzl, C., Shelton,

J.M., Richardson, J.A., *et al.* (2004b). Histone deacetylase 4 controls chondrocyte hypertrophy during skeletogenesis. Cell *119*, 555-566.

Verdel, A., and Khochbin, S. (1999). Identification of a new family of higher eukaryotic histone deacetylases. Coordinate expression of differentiation-dependent chromatin modifiers. J Biol Chem *274*, 2440-2445.

Verdin, E., and Ott, M. (2015). 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. Nat Rev Mol Cell Biol *16*, 258-264.

Villagra, A., Cheng, F., Wang, H.W., Suarez, I., Glozak, M., Maurin, M., Nguyen, D., Wright, K.L., Atadja, P.W., Bhalla, K., *et al.* (2009). The histone deacetylase HDAC11 regulates the expression of interleukin 10 and immune tolerance. Nat Immunol *10*, 92-100.

Wagner, G.R., and Payne, R.M. (2013). Widespread and enzyme-independent Nepsilon-acetylation and Nepsilon-succinylation of proteins in the chemical conditions of the mitochondrial matrix. J Biol Chem *288*, 29036-29045.

Wagner, S.A., Beli, P., Weinert, B.T., Nielsen, M.L., Cox, J., Mann, M., and Choudhary, C. (2011). A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. Mol Cell Proteomics *10*, M111 013284.

Wallingford, J.B., and Habas, R. (2005). The developmental biology of Dishevelled: an enigmatic protein governing cell fate and cell polarity. Development *132*, 4421-4436.

Wan, G., Fields, B.D., Spracklin, G., Shukla, A., Phillips, C.M., and Kennedy, S. (2018). Spatiotemporal regulation of liquid-like condensates in epigenetic inheritance. Nature *557*, 679-683.

Wang, A., Conicella, A.E., Schmidt, H.B., Martin, E.W., Rhoads, S.N., Reeb, A.N., Nourse, A., Ramirez Montero, D., Ryan, V.H., Rohatgi, R., *et al.* (2018a). A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. EMBO J *37*.

Wang, A.H., Kruhlak, M.J., Wu, J., Bertos, N.R., Vezmar, M., Posner, B.I., Bazett-Jones, D.P., and Yang, X.J. (2000). Regulation of histone deacetylase 4 by binding of 14-3-3 proteins. Mol Cell Biol *20*, 6904-6912.

Wang, H., and Ryu, W.S. (2010). Hepatitis B virus polymerase blocks pattern recognition receptor signaling via interaction with DDX3: implications for immune evasion. PLoS Pathog *6*, e1000986.

Wang, J., Choi, J.M., Holehouse, A.S., Lee, H.O., Zhang, X., Jahnel, M., Maharana, S., Lemaitre, R.,

Pozniakovsky, A., Drechsel, D.*, et al.* (2018b). A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. Cell *174*, 688-699 e616.

Wang, L., Lawrence, M.S., Wan, Y., Stojanov, P., Sougnez, C., Stevenson, K., Werner, L., Sivachenko, A., DeLuca, D.S., Zhang, L.*, et al.* (2011). SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. N Engl J Med *365*, 2497-2506.

Weber, S.C., and Brangwynne, C.P. (2012). Getting RNA and protein in phase. Cell *149*, 1188-1191.

Wei, W., Liu, X., Chen, J., Gao, S., Lu, L., Zhang, H., Ding, G., Wang, Z., Chen, Z., Shi, T.*, et al.* (2017). Class I histone deacetylases are major histone decrotonylases: evidence for critical and broad function of histone crotonylation in transcription. Cell Res *27*, 898-915.

Weinert, B.T., Iesmantavicius, V., Moustafa, T., Scholz, C., Wagner, S.A., Magnes, C., Zechner, R., and Choudhary, C. (2014). Acetylation dynamics and stoichiometry in Saccharomyces cerevisiae. Mol Syst Biol *10*, 716.

Weinert, B.T., Narita, T., Satpathy, S., Srinivasan, B., Hansen, B.K., Scholz, C., Hamilton, W.B., Zucconi, B.E., Wang, W.W., Liu, W.R.*, et al.* (2018). Time-Resolved Analysis Reveals Rapid Dynamics and Broad Scope of the CBP/p300 Acetylome. Cell *174*, 231-244 e212.

Wellen, K.E., Hatzivassiliou, G., Sachdeva, U.M., Bui, T.V., Cross, J.R., and Thompson, C.B. (2009). ATP-citrate lyase links cellular metabolism to histone acetylation. Science *324*, 1076-1080.

Wheeler, J.R., Matheny, T., Jain, S., Abrisch, R., and Parker, R. (2016). Distinct stages in stress granule assembly and disassembly. Elife *5*.

Wooderchak, W.L., Zang, T., Zhou, Z.S., Acuna, M., Tahara, S.M., and Hevel, J.M. (2008). Substrate profiling of PRMT1 reveals amino acid sequences that extend beyond the "RGG" paradigm. Biochemistry *47*, 9456-9466.

Woodruff, J.B., Ferreira Gomes, B., Widlund, P.O., Mahamid, J., Honigmann, A., and Hyman, A.A. (2017). The Centrosome Is a Selective Condensate that Nucleates Microtubules by Concentrating Tubulin. Cell *169*, 1066-1077 e1010.

Wu, D.W., Liu, W.S., Wang, J., Chen, C.Y., Cheng, Y.W., and Lee, H. (2011). Reduced p21(WAF1/CIP1) via alteration of p53-DDX3 pathway is associated with poor relapse-free survival in early-stage human papillomavirus-associated lung cancer. Clin Cancer Res *17*, 1895-1905.

Wu, J.Y., Xiang, S., Zhang, M., Fang, B., Huang, H., Kwon, O.K., Zhao, Y., Yang, Z., Bai, W., Bepler, G.*, et al.* (2018). Histone deacetylase 6 (HDAC6) deacetylates extracellular signal-regulated kinase 1 (ERK1) and thereby stimulates ERK1 activity. J Biol Chem *293*, 1976-1993.

Xiang, S., Kato, M., Wu, L.C., Lin, Y., Ding, M., Zhang, Y., Yu, Y., and McKnight, S.L. (2015). The LC Domain of hnRNPA2 Adopts Similar Conformations in Hydrogel Polymers, Liquid-like Droplets, and Nuclei. Cell *163*, 829-839.

Xie, Z., Zhang, D., Chung, D., Tang, Z., Huang, H., Dai, L., Qi, S., Li, J., Colak, G., Chen, Y.*, et al.* (2016). Metabolic Regulation of Gene Expression by Histone Lysine beta-Hydroxybutyrylation. Mol Cell *62*, 194-206.

Xiong, X., Panchenko, T., Yang, S., Zhao, S., Yan, P., Zhang, W., Xie, W., Li, Y., Zhao, Y., Allis, C.D.*, et al.* (2016). Selective recognition of histone crotonylation by double PHD fingers of MOZ and DPF2. Nat Chem Biol *12*, 1111-1118.

Xu, Z., Schaedel, L., Portran, D., Aguilar, A., Gaillard, J., Marinkovich, M.P., Thery, M., and Nachury, M.V. (2017). Microtubules acquire resistance from mechanical breakage through intralumenal acetylation. Science *356*, 328-332.

Yang, M.H., Laurent, G., Bause, A.S., Spang, R., German, N., Haigis, M.C., and Haigis, K.M. (2013). HDAC6 and SIRT2 regulate the acetylation state and oncogenic activity of mutant K-RAS. Mol Cancer Res *11*, 1072-1077.

Yang, M.H., Nickerson, S., Kim, E.T., Liot, C., Laurent, G., Spang, R., Philips, M.R., Shan, Y., Shaw, D.E., Bar-Sagi, D.*, et al.* (2012). Regulation of RAS oncogenicity by acetylation. Proc Natl Acad Sci U S A *109*, 10843-10848.

Yedavalli, V.S., Neuveut, C., Chi, Y.H., Kleiman, L., and Jeang, K.T. (2004). Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function. Cell *119*, 381-392.

Ying, Y., Wang, X.J., Vuong, C.K., Lin, C.H., Damianov, A., and Black, D.L. (2017). Splicing Activation by Rbfox Requires Self-Aggregation through Its Tyrosine-Rich Domain. Cell *170*, 312-323 e310.

Yoshizawa, T., Ali, R., Jiou, J., Fung, H.Y.J., Burke, K.A., Kim, S.J., Lin, Y., Peeples, W.B., Saltzberg, D., Soniat, M.*, et al.* (2018). Nuclear Import Receptor Inhibits Phase Separation of FUS through Binding to Multiple Sites. Cell *173*, 693-705 e622.

Yuan, H., Rossetto, D., Mellert, H., Dang, W., Srinivasan, M., Johnson, J., Hodawadekar, S., Ding, E.C.,

Speicher, K., Abshiru, N.*, et al.* (2012). MYST protein acetyltransferase activity requires active site lysine autoacetylation. EMBO J *31*, 58-70.

Zeng, L., Zhang, Q., Li, S., Plotnikov, A.N., Walsh, M.J., and Zhou, M.M. (2010). Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. Nature *466*, 258-262.

Zeng, M., Chen, X., Guan, D., Xu, J., Wu, H., Tong, P., and Zhang, M. (2018). Reconstituted Postsynaptic Density as a Molecular Platform for Understanding Synapse Formation and Plasticity. Cell *174*, 1172-1187 e1116.

Zeng, M., Shang, Y., Araki, Y., Guo, T., Huganir, R.L., and Zhang, M. (2016). Phase Transition in Postsynaptic Densities Underlies Formation of Synaptic Complexes and Synaptic Plasticity. Cell *166*, 1163-1175 e1112.

Zhang, C.L., McKinsey, T.A., Chang, S., Antos, C.L., Hill, J.A., and Olson, E.N. (2002). Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. Cell *110*, 479-488.

Zhang, G., Wang, Z., Du, Z., and Zhang, H. (2018a). mTOR Regulates Phase Separation of PGL Granules to Modulate Their Autophagic Degradation. Cell *174*, 1492-1506 e1422.

Zhang, H., Elbaum-Garfinkle, S., Langdon, E.M., Taylor, N., Occhipinti, P., Bridges, A.A., Brangwynne, C.P., and Gladfelter, A.S. (2015a). RNA Controls PolyQ Protein Phase Transitions. Mol Cell *60*, 220-230.

Zhang, J., Woodhead, G.J., Swaminathan, S.K., Noles, S.R., McQuinn, E.R., Pisarek, A.J., Stocker, A.M., Mutch, C.A., Funatsu, N., and Chenn, A. (2010). Cortical neural precursors inhibit their own differentiation via N-cadherin maintenance of beta-catenin signaling. Dev Cell *18*, 472-479.

Zhang, K., Daigle, J.G., Cunningham, K.M., Coyne, A.N., Ruan, K., Grima, J.C., Bowen, K.E., Wadhwa, H., Yang, P., Rigo, F.*, et al.* (2018b). Stress Granule Assembly Disrupts Nucleocytoplasmic Transport. Cell *173*, 958-971 e917.

Zhang, L., Liu, S., Liu, N., Zhang, Y., Liu, M., Li, D., Seto, E., Yao, T.P., Shui, W., and Zhou, J. (2015b). Proteomic identification and functional characterization of MYH9, Hsc70, and DNAJA1 as novel substrates of HDAC6 deacetylase activity. Protein Cell *6*, 42-54.

Zhang, M., Xiang, S., Joo, H.Y., Wang, L., Williams, K.A., Liu, W., Hu, C., Tong, D., Haakenson, J., Wang, C.*, et al.* (2014). HDAC6 deacetylates and ubiquitinates MSH2 to maintain proper levels of MutSalpha. Mol Cell *55*, 31-46.

Zhang, X., Yuan, Z., Zhang, Y., Yong, S., Salas-Burgos, A., Koomen, J., Olashaw, N., Parsons, J.T., Yang, X.J., Dent, S.R., *et al.* (2007). HDAC6 modulates cell motility by altering the acetylation level of cortactin. Mol Cell *27*, 197-213.

Zhang, Y., Gilquin, B., Khochbin, S., and Matthias, P. (2006). Two catalytic domains are required for protein deacetylation. J Biol Chem *281*, 2401-2404.

Zhang, Y., Li, N., Caron, C., Matthias, G., Hess, D., Khochbin, S., and Matthias, P. (2003). HDAC-6 interacts with and deacetylates tubulin and microtubules in vivo. EMBO J *22*, 1168-1179.

Zhang, Y., Yan, L., Zhou, Z., Yang, P., Tian, E., Zhang, K., Zhao, Y., Li, Z., Song, B., Han, J., *et al.* (2009). SEPA-1 mediates the specific recognition and degradation of P granule components by autophagy in C. elegans. Cell *136*, 308-321.

Zhao, D., Li, Y., Xiong, X., Chen, Z., and Li, H. (2017). YEATS Domain-A Histone Acylation Reader in Health and Disease. J Mol Biol *429*, 1994-2002.

Zhao, D.Y., Gish, G., Braunschweig, U., Li, Y., Ni, Z., Schmitges, F.W., Zhong, G., Liu, K., Li, W., Moffat, J., *et al.* (2016). SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination. Nature *529*, 48-53.

Zhou, X.Z., and Lu, K.P. (2016). The isomerase PIN1 controls numerous cancer-driving pathways and is a unique drug target. Nat Rev Cancer *16*, 463-478.

Acknowledgements

First and foremost, I would like to acknowledge my mentor Patrick Matthias for giving me the opportunity to perform my Ph.D. study in his lab and his constant mentoring. He gave me total freedom to pursue my interests in the laboratory, and thanks to him, I am confident that my Ph.D. work strongly reflected my philosophy about science. It is a precious thing for my future study that I could get this kind of training in early phase of my career. I also would like to express my gratitude to my external thesis committee members; Markus Affolter for his important advices and comments on my thesis, and Ivan Dikic for his insightful comments and inviting me as a speaker in his lab in the middle of my Ph.D. study.

I am deeply grateful to all collaborators who supported my Ph.D. work intellectually and technically. I would like to thank Daniel Hess for his constant contribution to to my project, and Jan Eglinger for his intellectual contribution on how to analyze images. I also would like to thank Chunaram Choudhary and Brian T. Weinert who are long-term collaborators of our laboratory and provided core datasets for my project. I would like to offer my special thanks to Anatol W. Fritsch, Moritz Kreysing and Anthony Hyman who provided me a chance to perform high-end microscopy at their institute on short notice and also technically contributed to the experiment. I cannot forget to thank Jeffrey Chao and his group members for generously providing me their resources, Luca Giorgetti for critical advice on my work with his expertise, and Syota Matsumoto and Nicolas Thoma for providing me with resources for protein purification from insect cells.

I would like to show my appreciation to all current and former lab members. Special thanks go to Longlong Wang and Shuang Song for their friendships in daily life beyond the lab. I would like to thank Etori Aguiar Moreira, Jacint Sanchez, Laura Sabou, and Jasmin Ostermayer for sharing time with me in the lab and fruitful discussion, Chun Cao and Gabriele Matthias for their constant support

to maintain working environment tidy and comfortable and caring so well. I would like to thank Roger G. Clerc for his kind advices on my career planning and stimulating collaboration work when he was taking a sabbatical in our lab. I also would like to thank all other former lab members: Yasuyuki Miyake, Amin M. Choukrallah, Anand Manoharan, Richard Heideman, Patricia Nigg, Fengyuan Tang, Vincent Pillonel and Daniel Weis.

I am particularly grateful for the assistance given by all FMI communities including the facilities such as microscopy and imaging, protein analysis, cell sorting, computational biology, protein structure, and transgenic mice facilities, as well as media kitchen. Without their persistent and excellent help, my Ph.D. work would not have been achieved. Honestly, I am really proud of spending over 5 years as a member of such a great institute. I want to thank the FMI Ph.D. program and the responsible person Elida Keller for her thoughtful support to my student life, and Piera Cicchetti for her kind advices on my student life. I would like to show my greatest appreciation to Nakajima Foundation in Japan, which partially funded me through my Ph.D. study and provided me opportunities to interact with other Japanese Ph.D. students also studying abroad.

My heartfelt appreciation goes to all friends I met during my staying in Basel. I want to thank all researcher friends in FMI, University of Basel, University of Zurich and ETH Zurich for sharing time at dinner/drinking parties to talk about science, life, future and so on. I also want to thank non-researcher colleagues living in Basel, including members of Japanischer-Verein Basel. Special thanks go to friends in Musik-Akademie Basel who vividly color my daily life with their excellent performance; their extreme dedication to music always inspires me to do science in such a lunatic way.

And last, I would like to thank my family for their constant support and encouragement through my entire life.